CONTENTS

Volume III

THE COMPLETE WORKS

of

JAMES ARTHUR HARRIS

Volume III

1918 - 1921

CONTENTS

Volume III

CONTENTS

Volume III

Contents, Volume III
Page 3.

Contents, Volume III
Page 4.

FURTHER ILLUSTRATIONS OF THE APPLICABILITY OF A COEFFICIENT MEASURING THE CORRELATION BE- TWEEN A VARIABLE AND THE DEVIATION OF A DEPENDENT VARIABLE FROM ITS PROBABLE VALUE

J. ARTHUR HARRIS

Station for Experimental Evolution, Cold Spring Harbor, New York

GENETICS

A Periodical Record of Investigations Bearing on
Heredity and Variation

FURTHER ILLUSTRATIONS OF THE APPLICABILITY OF A COEFFICIENT MEASURING THE CORRELATION BE-TWEEN A VARIABLE AND THE DEVIATION OF A DEPENDENT VARIABLE FROM ITS PROBABLE VALUE

J. ARTHUR HARRIS

Station for Experimental Evolution, Cold Spring Harbor, New York

FURTHER ILLUSTRATIONS OF THE APPLICABILITY OF A COEFFICIENT MEASURING THE CORRELATION BETWEEN A VARIABLE AND THE DEVIA- TION OF A DEPENDENT VARIABLE FROM ITS PROBABLE VALUE

J. ARTHUR HARRIS

Station for Experimental Evolution, Cold Spring Harbor, New York

TABLE OF CONTENTS

INTRODUCTORY

Eight years ago I pointed out (HARRIS 1909 a) that when y is some fraction of x the correlation between them, r_{xy}, while of descriptive value,

does not give all of the information which is required concerning the interrelationship of these two variables, and that a coefficient showing whether the value of y becomes relatively larger or smaller with increasing values of x, would have considerable analytical value.

I showed[1] then that if $z = px$, where $p = \bar{y}/\bar{x}$, the bars denoting population means,

$$r_{xz} = \frac{r_{xy} - v_x v_y \; \sqrt{x}/v_y}{\sqrt{1 - r_{xy}^2 + (r_{xy} - v_x/v_y)^2}}.$$

The purpose of this paper is to illustrate the range of usefulness of this coefficient by noting biological progress which has been made by its use, and by actually applying it to series of data which have not been heretofore fully analyzed by the higher statistical methods.

Before passing to actual illustrations of applicability of the coefficient, some questions of method should be considered.

In cases in which the coefficient r_{xz} is desired the correlation table has usually been formed to determine r_{vy}, where $v + y = x$. The mean and standard deviation for x made up of any number of components (HARRIS 1917 c) are well known (PEARL 1909, HARRIS 1918 a); when $x = v + y$

$$\bar{x} = \bar{v} + \bar{y}, \qquad \sigma_x = \sqrt{\sigma_v^2 + \sigma_y^2 + 2r_{vy}\sigma_v\sigma_y}$$

If one has made his summations about zero as origin, as suggested elsewhere (HARRIS 1910 b) it is quite easy to determine the necessary moments from those already available.

$$\bar{x} = \Sigma(x)/n = [\Sigma(v) + \Sigma(y)]/n$$
$$\sigma_x^2 = [\Sigma(v^2) + \Sigma(y^2) + 2\Sigma(vy)]/n - \bar{x}^2$$

The product moments are

$$\Sigma(xv) = \Sigma(v^2) + \Sigma(vy)$$
$$\Sigma(xy) = \Sigma(y^2) + \Sigma(vy)$$

$\Sigma(vy)$ is in each case the product moment already obtained in the calculation of r_{vy}.

In the original paper (HARRIS 1909 a), the method of determining the value of σ_z, which is essential for calculating the equation for the regression of z on x, was not indicated.

---

[1] The deduction of a usable formula for the problem then in hand was, as I have elsewhere pointed out, due to Professor PEARSON who, with characteristic generosity to his pupils, insisted that my name only appear in the paper.

The first and second moment of y for each grade of x, i.e., $\Sigma(y_x)$ and $\Sigma(y_x^2)$, are easily determined. The first of these will have been obtained in computing the population product moment, $S(xy)$, by the method cited above.

Remembering that $\Sigma(y_x)$ and $\Sigma(y_x^2)$ are the first and second moments about zero as origin for the several y arrays of x, the moments of z for each array are

$$\Sigma(z_x) = \Sigma(y_x) - n_x p x$$

From these the means of arrays which are required for plotting the regression curve are given at once by

$$\bar{z}_x = [\Sigma(y_x) - n_x p x]/n_x$$

For the whole population the first moment of z about o as origin is

$$S\Sigma(z_x) = S[\Sigma(y_x) - n_x p x] = 0$$

The second moments for the individual arrays are

$$\Sigma(z_x^2) = \Sigma(y_x^2) - 2px[\Sigma(y_x) - n_x p x] - n_x(px)^2,$$

and for the population

$$S\Sigma(z_x^2) = S\{\Sigma(y_x^2) - 2\ px[\Sigma(y_x) - n_x p x] - n_x(px)^2\},$$

or in some cases more conveniently for actual computation,

$$S(z^2) = S(y^2) - 2pS(xy) + S[n_x(px)^2],$$

where S denotes summation of values of y arrays of x, or throughout the population. Thus

$$\sigma_z = [S(z^2)/N]^{\frac{1}{2}}$$

It is quite possible to determine the correlation between x and z, the deviation of y from its probable value, directly.

Remembering that $[\Sigma(y_x) - n_x p x]$ is the first moment about zero as origin of z for any array, the product moment for the population is

$$S x[\Sigma(y_x) - n_x p x] ,$$

where S denotes summation of the values of the y arrays of x.

Practically it is more convenient to determine the product moment from

$$S(xz) = S(xy) - pS(x^2),$$

where $S(xy)$ and $S(x^2)$ are the product moments of x and y and the second moment of x for the population.

The method has been most extensively applied to problems of fertility and fecundity. Thus the relationship between the number of ovaries formed and the number of ovaries developing into fruits has been investigated in the inflorescence of Staphylea (HARRIS 1909), Celastrus (HARRIS 1909 b) and Crinum (HARRIS 1912). In Staphylea and Crinum inflorescences which produce larger numbers of flowers mature relatively fewer fruits. In Celastrus there is apparently no relationship between the number of flowers formed and the capacity of the inflorescence for maturing the ovaries into fruits.

In the fruit, the relationship between the total number of ovules laid down and the deviation of the number of seeds matured from their probable number has been investigated in Sanguinaria (HARRIS 1910 a).

For *Phaseolus vulgaris* a first study (HARRIS 1913) of 53 series comprising 166,130 pods and a supplemental investigation of 16 series comprising 56,698 pods (HARRIS 1917 a) leave no doubt that the pods with the larger number of ovules mature relatively fewer of their ovules into seeds. The same relationship holds in the arborescent legume, *Cercis canadensis,* as is shown by studies based on massed data (HARRIS 1914 a) and on series from individual trees (HARRIS 1914 b).

The relationship found in Cercis and Phaseolus is not universal for the Leguminosae. In a series of 1427 pods of Robinia (HARRIS 1909 a), the pods with larger numbers of ovules mature a relatively higher proportion of their ovules into seeds. The correlation between the actual number of ovules formed and the actual number of seeds developing is $r_{os} = .693 \pm .009$, while that between the number of ovules formed and the deviation of the number of seeds matured from their probable value is $r_{oz} = .365 \pm .015$.

That this result represents a real biological relationship is indicated by the correlations, hitherto unpublished, for the individual trees. Only three of the twelve constants in table 1 are negative in sign. No one of these can be regarded as statistically significant when the probable error is taken into consideration, while seven of the nine positive coefficients must be looked upon as statistically trustworthy.

The formula has also been advantageously applied to the problem of the interrelationship of the number of male and female flowers in the inflorescence of the aroid Arisarum (HARRIS 1916 a) and that of the interdependence of numbers of stamens and pistils in the ranunculaceous genus, Ficaria (HARRIS 1918). In Arisarum the relative number

TABLE I

Relationship between seed and ovule number in Robinia.

Number of tree	Number of pods	r_{os}	r_{oz}	$r_{oz}/E_{r_{oz}}$
1	122	.524±.044	—.055±.061	(—)0.90
2	64	.802±.030	.468±.066	7.09
3	111	.478±.049	.074±.064	1.16
4	102	.430±.054	.105±.066	1.59
5	122	.671±.034	.338±.054	6.26
6	120	.533±.044	.272±.057	4.77
7	120	.259±.057	—.115±.061	(—)1.89
8	159	.614±.033	.216±.051	4.24
9	128	.590±.040	.350±.052	6.73
10	78	.507±.057	.225±.073	3.08
11	105	.714±.032	—.044±.066	(—)0.67
12	196	.797±.018	.489±.037	13.22
All trees	1427	.693±.009	.365±.015	24.33

of pistillate flowers increases as the total number of flowers per inflorescence increases. In Ficaria the relative number of pistils increases as the total number of sporophylls becomes larger.

Dr. BLAKESLEE and I (1918) have applied this coefficient to the determination of the relationship between the total annual egg production and the monthly egg production of White Leghorn fowl. We have there shown by means of this coefficient that the winter months, November, December, January and February, and the following autumn months, August, September and October, show an increase over their theoretical quota of eggs when the annual total egg production rises above the normal. That is, r_{ez}, the correlation between total annual egg production and the deviation of the monthly production from its probable value, is on the whole significantly and substantially positive. The spring and summer months, April, May, June and July, show negative values of r_{ez}, that is, they make a lower relative contribution to the annual total than might be expected when the total varies in the direction of an increase above the normal egg production of the flock as a whole.

FURTHER ILLUSTRATIONS

Illustration 1. Proportionality of parts in Paramecium

JENNINGS (1911) in his masterly investigation of assortative mating in Paramecium, has given data for determining the relationships between

(a) distance from the anterior end of the organism to the posterior margin of the mouth, (b) distance from the posterior margin of the mouth to the posterior end of the organism, and (c) the total length in series of conjugant and non-conjugant Paramecia.

He has calculated and discussed for a purpose which does not concern us here the correlations between certain of these dimensions. All the correlations between the absolute measurements, calculated from his data, are given in table 2.

TABLE 2

Relationship between total length, l, and anterior length, a, and between anterior length and posterior length, p, in Paramecium.

Series*	r_{ap}	r_{la}	r_{lz}	$r_{lz}/E_{r_{lz}}$
Lot 7, table 40	.246±.050	.735±.025	—.733±.025	3.0
Lot 7, table 41	.570±.047	.893±.014	—.427±.057	7.5
Lot 19, table 51	.382±.044	.832±.016	—.434±.042	10.4
Lot 19, table 52	.671±.034	.939±.074	—.196±.060	3.3
Lot 22, table 55	—.403±.026	.485±.049	—.282±.060	4.7
Lot 22, table 56	.620±.025	.906±.073	—.405±.034	12.0
Lot 24, table 64	.277±.040	.820±.014	—.282±.040	7.1
Lot 24, table 65	.488±.031	.885±.089	—.288±.038	7.6

* Lot 7, table 40, conjugants of wild cultures. 2. Lot 7, table 41, non-conjugants of wild cultures. 3. Lot 19, table 51, conjugants of race *g*. 4. Lot 19, table 52, non-conjugants of race *g*. 5. Lot 22, table 55, wild culture conjugants not yet separated. 6. Lot 22, table 56, wild culture conjugants about twelve hours after separation. 7. Lot 24, table 64, unseparated conjugants of race *k*. 8. Lot 24, table 65, conjugants of race *k* about twelve hours after separation.

In this table the constants for anterior and posterior length are arranged in pairs of conjugants and non-conjugants or ex-conjugants. In every instance the correlation between the anterior and posterior portions of conjugants is lower than that between the same dimensions in non-conjugants or ex-conjugants.[2]

All of these values are low, as JENNINGS has noted. In the case of lot 22 the coefficient for the conjugants is actually negative in sign.

The correlation between total length and the length of the section anterior to the mouth is high. In every case the value of r_{la} is higher in non-conjugants or ex-conjugants than in conjugants. The additional relationship to be brought out by the formulae here under discussion is

[2] See in this connection the discussion by JENNINGS (1911, pp. 65-66, 71-73).

that between the total length of the organism and the relative length of either anterior or posterior element.

JENNINGS (1911, p. 63) has emphasized the high variability of the post-oral dimension. Table 3, in which all the coefficients of variation

TABLE 3

Coefficients of variation for anterior and posterior fractions of length in Paramecium.

Series	Total length	Anterior length	Posterior length	$V_p - V_a$
Lot 7, table 40	6.90	5.79	13.86	+ 8.07
Lot 7, table 41	9.68	8.76	14.53	+ 5.77
Lot 19, table 51	8.55	7.78	14.31	+ 6.53
Lot 19, table 52	12.18	12.08	15.13	+ 3.05
Lot 22, table 55	5.48	7.38	14.29	+ 6.91
Lot 22, table 56	7.96	7.28	11.20	+ 3.92
Lot 24, table 64	6.34	6.42	9.98	+ 3.56
Lot 24, table 65	6.62	6.45	9.35	+ 2.90

are laid side by side, fully confirms his conclusion in this regard. Utilizing these coefficients of variation we obtain the values for the correlation between total length and the deviation of the anterior length from its probable value, given in the fourth column of table 2.

These constants are negative in sign throughout, and while variable in magnitude all may reasonably be considered statistically significant in comparison with their probable errors.

Thus when Paramecium varies in length both anterior and posterior fractions of the body contribute to this variation, but as length increases the anterior portion becomes relatively shorter.

For one series, the unseparated conjugants of race k, I have determined the regression of the anterior length on total length and the regression of the deviation of the anterior length from its probable value on the total length of the organism. The equations are

$$a = 3.0359 + .4977\ l$$
$$\sigma_x = .7616, \quad z_a = 3.3743 - .1135\ l$$

The equations and empirical means are represented graphically in diagrams 1 and 2. In both cases the relationships are sensibly linear.

DIAGRAM 1.—Relationship of anterior length to total length in Paramecium. Compare diagram 2.

DIAGRAM 2.—Relationship of deviation of anterior length from its probable value to total length in Paramecium. Compare diagram 1.

Illustration 2.   Absence of relationship between size of litter and sex in swine

PARKER and BULLARD (1913) have discussed the possible relationship between the size of the litter and sex in the contents of 1000 uteri of swine. From a simple percentage table they state that the relative numbers of males and females are "even in the extreme cases so nearly uniform that we may conclude with reasonable assurance that there is no intimate relation between sex and the size of the litters."

The correlations between the total numbers of pigs in the litter, l, and the number of males, m, and females, f, may be deduced from their data. They are

$$r_{lm} = .6833 \pm .0114$$
$$r_{lf} = .6875 \pm .0112$$

From these and the three coefficients of variation one may deduce

For males $r_{lz} = - .0177 \pm .0213$
For females $r_{lz} = + .0177 \pm .0213$

The correlation is sensibly zero, with regard to its probable error. This method of analysis therefore fully confirms the conclusion drawn by PARKER and BULLARD.

Illustration 3.   Proportion of pistillate and hermaphrodite flowers in the inflorescence of the composite Homogyne

LUDWIG (1901) has given data for the correlation between the number of pistillate and the number of hermaphrodite flowers in the inflorescence of Homogyne. From his data we deduce

For hermaphrodite flowers, h,

$$\bar{h} = 31.8333, \; \sigma_h = 7.3981, \; V_h = 23.240$$

For pistillate flowers, p,

$$\bar{p} = 10.5370, \; \sigma_p = 2.6460, \; V_p = 25.112$$

For total flowers, f,

$$\bar{f} = 42.3704, \; \sigma_f = 8.7749, \; V_f = 20.7099$$

For hermaphrodite and pistillate flowers,
$$r_{hp} = .3899 \pm .0449$$

For total flowers and hermaphrodite flowers,
$$r_{fh} = .9607 \pm .0049, \; r_{fz_h} = .2429 \pm .0499$$

For total flowers and pistillate flowers,
$$r_{fp} = .6303 \pm .0339, \; r_{fz_p} = - .2429 \pm .0499$$

It follows, therefore, that in the larger heads the purely pistillate flowers are relatively less, and the hermaphrodite flowers relatively more, numerous.

Illustration 4. Fertility of capsules and viability of seed in carnation crosses

STUART (1912) has recorded the number of seeds obtained and the number germinated, planted into the field, and producing flowers in various carnation crosses. Our problem is to determine whether the seeds which come from capsules producing a large number of seeds are relatively more (or less) viable than those from capsules producing small numbers.

Using his two larger tables of data, tables 3 and 6,[3] and confining attention to the relationship between number of seeds per capsule, and the number which germinated, I find

For commercial $\times$ commercial, STUART's table 3, $N = 23$,
$$r_{sg} = .775 \pm .056, \qquad r_{sz} = -.072 \pm .141$$
For single flower $\times$ double flower, STUART's table 6, $N = 32$,
$$r_{sg} = .649 \pm .069, \qquad r_{sz} = -.118 \pm .118.$$

The signs are both negative, indicating a relatively higher failure to germinate among the seeds which are produced many in a capsule. With regard to their probable errors, the constants are untrustworthy. Because so few observations are available, no biological significance is attached to these two series, which serve merely as another illustration of the kind of problems to which the method may be applied.

Illustration 5. Relationship between the total number of pedicels and the number of abnormal pedicels in Spiraea Vanhouttei

In Spiraea Vanhouttei the pedicels of the umbel-like raceme normally produce but a single flower each. An abnormal condition in which one or more pedicels may bear a relatively large number of flowers is frequently observed (HARRIS 1917 d).

Let x be the total pedicels in an inflorescence and a the number which are abnormal. Then if abnormality be distributed purely at random among the pedicels one would expect material values of r_{xa}. The correlation r_{xz} meets our requirements since it shows whether inflorescences with a large number of rays have relatively more or fewer of their rays abnormal than those with a small number.

---

[3] In table 6 the cases in which the seeds are not normally developed are omitted.

During the last fifty years a great deal has been said about the influence of nutrition, vegetative vigor, etc., upon the development of anomalies. If a larger number of rays indicates greater vigor or better nutrition one might *a priori* expect larger inflorescences to have a proportionately higher number of branched rays, providing of course, that the classic theories are true.

The constants for a short series of data collected in 1906 were published in 1909. Since then a large number of determinations have been made on a general sample of inflorescences from a number of shrubs in 1909 and from three large individual shrubs in 1913.

In the latter series the data have been analyzed in two ways. First, the inflorescences which contain at least a single abnormal pedicel have been used as the basis of the correlations. These are designated as the abnormal inflorescences. Second, the normal inflorescences from the same plants have been included and counted as zero in the distribution of number of abnormal rays.

The results are:

For 1906[4] $r_{xa} = + .121 \pm .034$
$$r_{xz} = - .071 \pm .034$$

For 1909. Massed statistics. Inflorescences producing some abnormal pedicels ($N = 785$),

$$r_{xa} = + .1542 \pm .0235$$
$$r_{xz} = - .0915 \pm .0239$$

For 1909. Massed statistics. All 2040 inflorescences,

$$r_{xa} = + .1584 \pm .0146$$
$$r_{xz} = + .0370 \pm .0149.$$

For 1913. Individual plants. Inflorescences producing some abnormal pedicels,

Plant 1. $N = 747$ inflorescences.

$$r_{xa} = + .0880 \pm .0244$$
$$r_{xz} = - .2846 \pm .0227$$

Plant 2. $N = 641$ inflorescences.

$$r_{xa} = + .1148 \pm .0263$$
$$r_{xz} = - .3855 \pm .0227$$

---

[4] Since in the 1906 series only synanthies were observed, the total pedicels (including those with synanthies) were counted and the correlation is between the total pedicels and the pedicels with synanthies. In the 1909 material in which synanthies were rare, the correlation was determined for total pedicels and abnormal pedicels, abnormality being synanthy or any degree of branching.

Plant 3. $N = 548$ inflorescences.

$$r_{xa} = + .0941 \pm .0285$$
$$r_{xz} = - .2849 \pm .0265$$

For 1913.   Individual plants.   All inflorescences,
Plant 1. $N = 1135$ inflorescences.

$$r_{xa} = - .0067 \pm .0200$$
$$r_{xz} = - .2125 \pm .0191$$

Plant 2. $N = 975$ inflorescences.

$$r_{xa} = - .0821 \pm .0214$$
$$r_{xz} = - .3681 \pm .0187$$

Plant 3. $N = 912$ inflorescences.

$$r_{xa} = -.0360 \pm .0223$$
$$r_{xz} = - .2342 \pm .0211$$

For all the samples of inflorescences in which there is at least one abnormal pedicel the correlations between the total number of pedicels and the number of normal pedicels is positive in sign and perhaps statistically significant, but low in actual magnitude.   Thus the number of abnormal pedicels increases on the average as the total number of pedicels per inflorescence becomes larger.   The relationships are, however, very slight indeed.

For these five series the correlation between the total number of pedicels and the deviation of the abnormal pedicels from their probable value, is negative in sign.   Thus the larger inflorescences have a relatively smaller proportion of abnormal pedicels than do those with a smaller total number of pedicels.

In the four series in which the wholly normal inflorescences are included, the correlations between total number of pedicels and number of abnormal pedicels is positive in 1909 but negative throughout and insignificant in magnitude in 1913.   The three series from individual shrubs studied in 1913 show low but significantly negative correlations between the total number of pedicels per inflorescence and the deviation of the number of abnormal inflorescences from their probable value.   The constant for the heterogeneous data of 1909 is positive but insignificant.

Taking the data altogether, there can be no reasonable doubt that the relative number of abnormal pedicels decreases as the total number of abnormal pedicels increases.

This is shown in diagram 3, which represents the regression of the deviation of the number of abnormal rays from their probable value on

the total number of rays in the series showing the lowest, and in one of
these showing the highest, correlation.

DIAGRAM 3.—Regression of the deviation of the number of abnormal pedicels from
their probable value on the total number of pedicels in Spiraea.

The standard deviations are:
    For 1909, $\sigma_z = 2.4691$.
    For 1913, plant 1, $\sigma_z = 3.0527$.
The regression equations are:
    For 1909, $z = 1.3186 - .0609\ x$.
    For 1913, $z = 3.5036 - .2212\ x$.

*Illustration 6. Interrelationship of cotyledons and primordial leaves in a
race of Phaseolus vulgaris highly variable in seedling characters*

The ontogeny of the flowering plant is usually divided into a primary
period of dependent development and a sharply separated period of inde-
pendent growth. The leaf homologs which are developed during the first
period are usually of a very definite form and number, and frequently

highly differentiated from those developed after growth is resumed as an independent organism.

Those rare cases in which the number of leaf homologs which are laid down during the development of the seed is highly variable seem to offer especially favorable opportunities for the morphologist to learn something of the interrelationship of the two forms which may be assumed by homologous organs.

Suitable material for such work is furnished by a race of *Phaseolus vulgaris* (HARRIS 1916 b) having a modal number of four cotyledons and four primordial leaves, but highly variable in the number of both of these organs.

The correlations and regression equations are :[5]

Correlation between number of cotyledons and number of leaves:

Less mature series, $r = .1170 \pm .0320$.
$$c = 3.4759 + 0.0421\ l$$
$$l = 3.2919 + 0.3252\ c$$
More mature series, $r = .1568 \pm .0236$.
$$c = 3.5128 + 0.0499\ l$$
$$l = 3.0039 + 0.4922\ c$$
Combined series, $r = .1386 \pm .0193$.
$$c = 3.5030 + 0.0464\ l$$
$$l = 3.3565 + 0.4142\ c$$

The correlations between number of cotyledons and number of primordial leaves indicate slender interrelationships in this newly originated race between characters which are usually thought of as highly correlated.

Note from the equations that there is on the average a change of .3 to .5 of a leaf for each variation of one cotyledon, but that there is a change of only .04 of a cotyledon for a variation of one leaf.

The results for the combined series are represented graphically in diagram 4. Regression is for each character very nearly linear.

Correlation between total leaf homologs and number of cotyledons :[6]

Less mature series, $r = .4328 \pm .0264$.
$$c = 2.4595 + 0.1413\ h$$

---

[5] The plants were graded in a less mature and a more mature series as a precaution against recording as primordial leaves those really developed subsequently to the maturation of the seed.

[6] The following correlations are published here for the first time.

DIAGRAM 4.—Relationship between number of cotyledons and number of primordial leaves in a tetracotyledonous race of *Phaseolus vulgaris*.

More mature series, $r = .4337 \pm .0197$.
$$c = 2.6704 + 0.1260.h$$
Combined series, $r = .4312 \pm .0084$.
$$c = 2.5954 + 0.1314\ h$$

The straight line showing the regression of number of cotyledons, c, on total leaf homologs, h, is shown with the empirical means in diagram 5.

Correlation between total leaf homologs and number of leaves:

Less mature series, $r = .9459 \pm .0034$.
$$l = -\ 2.4595 + 0.8586\ h$$

DIAGRAM 5.—Relationship between number of cotyledons and number of leaves and total number of leaf homologs (cotyledons + leaves) in Phaseolus.

More mature series, $r = .9579 \pm .0020$
$$l = -\ 2.6704 + 0.8740\ h$$
Combined series, $r = .9533 \pm .0018$.
$$l = -\ 2.5953 + 0.8685\ h$$

Diagram 5 gives the regression of number of leaves on the number of leaf homologs (cotyledons + leaves). The solid dots representing the empirical mean leaf number lie practically on the theoretical line.

Correlation between total leaf homologs and the deviation of the number of leaves from their probable value:

Less mature series, $r = .6936 \pm .0226$.
More mature series, $r = .7642 \pm .0185$.
Combined series, $r = .7378 \pm .0089$.

The correlations between total leaf homologs and the deviations of the cotyledons from their probable value are numerically identical with the foregoing but negative in sign.

The results show a high degree of consistency of the two series. These final constants show that when the total number of leaf homologs increases, the variation is due to a far greater extent to the laying down of a greater number of primordial leaves than to the formation of a larger number of cotyledons. The regression equations for both the deviation of the number of leaves from their probable value (z_l) and the deviation of the number of cotyledons from their probable value (z_c) are given by

$$\sigma_z = .874557$$
$$z_c = 2.5954 - 0.3007\ h$$
$$z_l = -\ 2.5954 + 0.3007\ h$$

and represented in diagram 6. The results are clearly linear.

Illustration 7. Change in proportion of parts in developing trout

JENKINSON (1912) has given data for total length and length of head for three growth stages in the American rainbow trout. His constants are:

	Total length, l CV	Head length, h CV	Total length and head length r_{lh}
Stage 1	13.11	19.68	.945
Stage 3	5.11	6.57	.729
Stage 5	7.74	6.55	.848

DIAGRAM 6.—Regression of the deviation of the number of cotyledons and of the number of leaves from their probable numbers on the total number of leaf homologs in Phaseolus. Note that the theoretical lines and the empirical means are identical but opposite in sign.

Whence we deduce:

$$\text{Stage } 1, \; r_{lz} = + .649 \pm .027.$$
$$\text{Stage } 3, \; r_{lz} = - .071 \pm .047.$$
$$\text{Stage } 5, \; r_{lz} = - .534 \pm .038.$$

These results bring out especially well the analytical value of the method here proposed. The correlations between total length and head length are *roughly* the same in the three stages of development. The correlation between total length and deviation of head length is at first strongly positive, then sensibly zero in the third stage and finally becomes negative and of the order — .5 in the fifth stage.

Thus while in each stage the larger individuals have the larger heads, the relationships between the total length of the body and the proportional length of the head changes greatly during development.

Illustration 8.   Relationship between total solids and sucrose content in the juice of sugar beets

Sucrose content of the juice of the sugar beet has been one of the classic examples of variation in text-books on genetics.   Nevertheless our knowledge of the problems are, as indicated by a recent review .(HARRIS 1917 b), far from complete.   Relatively little is known, for example, of the correlation between total solids and sucrose content. Yet this problem is not merely of physiological and genetical interest but of very great practical importance as well, since the coefficient of purity of the juice is an important factor in sugar manufacture.

Table 4 shows the correlation between total solids and sucrose content in 475 Nevada sugar beets.   Nominally this series is composed of a number of commercial varieties, but since it has been shown elsewhere (HARRIS and GORTNER 1913) that the correlation between weight and composition of juice is not greatly influenced by the combination of these nominal varieties, such heterogeneity as exists probably does not influence materially the correlations to be deduced.   For purposes of comparison the constants for a series of 61 determinations from Washington are added.

The Nevada series gives:
  For total solids,
          Mean $= 18.021$, $\sigma = 3.7224$, CV $= 20.656$
  For sucrose,
          Mean $= 14.989$, $\sigma = 4.0727$, CV $= 27.171$
  For total solids and sucrose,
                    $r = .882 \pm .007$
  For total solids and deviation of sucrose from its probable value,
                  $r_{sz} = + .251 \pm .029$

For the Washington series the results are:
  For total solids,
          Mean $= 16.445$, $\sigma = 1.5896$, CV $= 9.666$
  For sucrose,
          Mean $= 11.838$, $\sigma = 1.9098$, CV $= 16.133$
  For total solids and sucrose,
                    $r = .866 \pm .022$

TABLE 4

Total solids and sucrose content in sugar beets.

Total solids	Sucrose content																									Totals
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
6			2	1																						3
7																										
8	1	1					1																			1
9				1			1																			3
10						1																				1
11						3	4	2																		7
12				1	1		3	5	2		2															14
13						1	1	4	3	10	8	1														21
14		1					1	3	8	6	15	6	1				1									36
15						1	1		7	3	6	9	5	4	1											41
16							1	1		1	1	9	13	16	5	1	1									37
17				1	1			1		1	2	5	9	10	20	9	6									40
18								1			2	3	7	4	7	17	28	3		1						53
19								1		1	1	1	3	1	4	18	8	5		1						42
20												1		4	2	2	5	9	10	3	1					58
21									1			2		3	1	3	1	1	5	8	2					34
22												1				2	1	1	5	9	5	2				31
23																1	1		3	1	3		2		1	20
24																2	1	1		9	5	2	1			22
25																1			3	1	3					4
26																					2	2				4
27																				1						2
34																										1
Totals	1	2	2	4	2	6	13	18	21	22	37	38	38	42	40	53	51	19	23	24	13	4	1		1	475

For total solids and deviation of sucrose from its probable value,

$$r_{sz} = .471 \pm .067$$

Thus in both cases the correlation between total solids and sucrose is moderately high, as is necessarily the case because of the fact that total solids are to a considerable extent made up of sucrose.

Quantity of sucrose is in both cases more variable than the total amount of solids. The correlation between total solids and the relative amount of sucrose is in both cases positive and of a substantial order of magnitude. Beet juice with higher total solids contains both absolutely more and relatively more sugar than that of beets with low total solid content.

Another way of approaching the problem is to calculate the correlations between total solids and the coefficient of purity, which is often determined.

Nevada series,

$$r_{sp} = .340 \pm .027$$

Washington series,

$$r_{sp} = .529 \pm .062$$

These values are also positive and substantial, but their interpretation involves difficulties due to spurious correlation.

Illustration 9.   Relationship between total number of spikelets and number of sterile spikelets in wheat

GRANTHAM and GROFF (1916) have given data for the correlation between the total number of spikelets and the number of sterile spikelets at the base of the head in varieties of wheat. Let t = total spikelets and s = sterile spikelets. Then from their two tables I find, without applying SHEPPARD'S correction,

Bearded varieties,

$$r_{ts} = .635 \pm .039, \; r_{ts_s} = + .164 \pm .063$$

Beardless varieties,

$$r_{ts} = .542 \pm .054, \; r_{ts_s} = + .232 \pm .072$$

For both groups of varieties there is a positive correlation between total spikelets and sterile spikelets and between total spikelets and the deviation of the number of sterile spikelets from their probable value. Sterility is, therefore, not merely absolutely but relatively more frequent in varieties with larger numbers of spikelets.

Illustration 10.   Viability of dominants and recessives in F_2 generation of Mendelian hybrids

F_1 D(R) $\times$ D(R), according to Mendelian theory, gives in F_2 a population with somas 3 D : R.   Let there be N individual families in this population and the number of individuals per family be variable.   This variation may be due solely to fluctuations in the number of zygotes produced, or it may be in part attributable to the failure of some of the zygotes to reach maturity.   This failure may be random or differential with respect to the alternative characters involved.   If differential, one might expect a correlation between the actual number of individuals per family and the deviation from the probable number of either of the alternative types of individuals which it contains.

In current Mendelian literature discrepancies between observed and theoretical ratios are often explained as due to selective fertilization or to a lower viability of particular zygotes.

If the latter explanation be correct r_{lz}, the correlation between the number in the litter and the deviation from their probable frequency of the number which shows a particular character, should have a statistically significant value.   As an illustration of the kind of data to which this test may be applied, I take YULE's (1914) table of DARBISHIRE's (1904) results for albinos in mice families.   If the fertilized ovum which is to produce an albino be less capable of development than that which is to develop into a normal individual r_{lz} should take a substantial positive value, i.e., the relative numbers of albino young should be larger in the large litters, since the small litters are assumed to be small *in part* because albino-producing fertilized ova are supposed to be less viable.

The actual results are :

$$\bar{l} = 4.5868 \qquad \bar{a} = 1.1322$$
$$\sigma_l = 1.7518 \qquad \sigma_a = 1.0752$$
$$V_l = 38.193 \qquad V_a = 94.963$$

$$r_{la} = .3756 \pm .0527$$
$$r_{lz} = -.0287 \pm .0612$$

where l = number in litter, a = number of albinos in litter, and z = deviation of number of albinos from their probable value.

The constant of critical significance, r_{lz}, is sensibly zero, with due regard to its probable error.   Hence there is no evidence of a greater intra-uterine mortality of albino zygotes.

RECAPITULATION

In many instances the biologist has to consider the relationship between a measurement and some of its logical subdivisions or components. The first has been called an independent and the second, which is always some fraction of the first, a dependent variable.

For the analysis of such relationships two coefficients are required, the correlation between the variable and the dependent variable and the correlation between the variable and the deviation of the dependent variable from its probable value on the assumption that the relative magnitude of the dependent variable is independent of the magnitude of the variable.

This paper gives (a) the supplementary formulae which are required in certain cases in which the correlation between a variable and the deviation of a dependent variable from its probable value is to be computed, and (b) a series of illustrations of the applicability of this coefficient, drawn from a wide range of biological phenomena.

The method has been most extensively applied to the problems of the physiology of seed production in plants. The relationship of the number of ovaries which develop to maturity to the total number of ovaries formed, and the relationship of the number of seeds which mature to the total number of ovules laid down, have been determined in a considerable number of plant forms. That it may be useful in the study of the relationship of seed viability to the number of seed formed is shown by analysis of meagre data for carnation crosses.

The formulae have been applied to the problem of sex by an investigation of the relative proportion of macrosporophylls and microsporophylls in data for Arisarum, Ficaria and Homogyne, and for the relative numbers of males and females in litters of swine.

Under certain conditions the method may be of use in testing the assumption of the existence of a differential viability in Mendelian dominants and recessives.

The usefulness of the method in morphology has been illustrated by its application to the problem of the relationship between total length and anterior length in Paramecium, to that of the relative size of the head in developing trout, to that of the relative frequency of abnormal pedicels in the inflorescence of Spiraea, and to that of the relationship of the numbers of cotyledons and primordial leaves to the total number of leaf homologs in a highly variable race of Phaseolus. Such studies have not merely an independent morphological value, but have their bearing

upon such biological theories as DRIESCH's annunciation that in a differentiated system the proportionality of the parts is absolutely independent of the size of the system.

In dealing with materials of economic importance, the method has been applied to the problem of the relative proportion of sterile spikelets in varieties of wheat with varying numbers of total spikelets per head, to the relationship between total solids and sucrose content in the juice of sugar beets, and has been found of the greatest value in analyzing the relationship between the egg production of the individual months and that of the entire year in the domestic fowl.

These illustrations are perhaps sufficient to show that the coefficient has great analytical value, and should have the widest usefulness.

Finally it must be understood that the conclusions drawn from the illustrations given in this paper are not in all cases to be extended beyond the specific series of data to which the formulae have been applied. The data employed are, *in some instances*, intended to illustrate the type of biological problem to which the formulae may be applied when more extensive data are available. No attempt has been made to discuss in detail the biological significance of results of so diverse kinds of illustrations. Such discussion must be left to the specialist in the particular field in which the observations fall.

LITERATURE CITED

DARBISHIRE, A. D., 1904 On the result of crossing Japanese waltzing with albino mice. Biometrika **3**: 1-51.

GRANTHAM, A. E., and GROFF, F., 1916 Occurrence of sterile spikelets in wheat. Jour. Agr. Res. **6**: 235-250.

HARRIS, J. ARTHUR, 1909 a The correlation between a variable and the deviation of a dependent variable from its probable value. Biometrika **6**: 436-443.

1909 b Correlation in the inflorescence of *Celastrus scandens*. Ann. Rep. Mo. Bot. Gard. **20**: 116-122.

1910 a A quantitative study of the morphology of the blood root, *Sanguinaria canadense*. Biometrika **7**: 305-351.

1910 b The arithmetic of the product moment method of calculating the coefficient of correlation. Amer. Nat. **44**: 693-699.

1912 Biometric data on the inflorescence and fruit of *Crinum longifolium*. Ann. Rep. Mo. Bot. Gard. **23**: 75-99.

1913 On the relationship between the number of ovules formed and the capacity of the ovary for maturing its ovules into seeds. Bull. Torr. Bot. Club **40**: 447-455.

1914 a On the relationship between the number of ovules formed and the number of seeds developing in Cercis. Bull. Torr. Bot. Club. **41**: 243-256.

1914 b Further observations on the relationship between the number of ovules formed and the number of seeds developing in Cercis. Bull. Torr. Bot. Club **41**: 533-549.

HARRIS, J. ARTHUR, 1916 a On the distribution and correlation of the sexes (staminate and pistillate flowers) in the inflorescence of the aroids *Arisarum vulgare* and *Arisarum proboscidium*. Bull. Torr. Bot. Club **42**: 663-673.

1916 b A tetracotyledonous race of *Phaseolus vulgaris*. Mem. N. Y. Bot. Gard. **6**: 229-244, *figs. 1-3*.

1917 a Supplementary determinations of the relationship between the number of ovules per pod and fertility in Phaseolus. Genetics **2**: 282-290.

1917 b Biometric studies on the somatic and genetic physiology of the sugar beet. Amer. Nat. **51**: 507-512.

1917 c The correlation between a component, and between the sum of two or more components, and the sum of the remaining components of a variable. Quart. Publ. Amer. Stat. Ass. **15**: 854-859. *pls. 30, 31*.

1918 a The interrelationship of the number of stamens and pistils in the flowers of Ficaria. Biol. Bull. **34**: 7-17, *figs. 1-5*.

1918 b On the distribution of abnormalities in the inflorescence of *Spiraea Vanhouttei*. Amer. Jour. Bot. **4**: 624-636, *diagrams 1-4, pls. 30, 31*.

HARRIS, J. ARTHUR, and BLAKESLEE, A. F., with the coöperation of KIRKPATRICK, WM. F., 1917 Inter-periodic correlation in the egg production of the domestic fowl. Proc. Nat. Acad. Sci. **3**: 565-569, *figs. 1, 2*.

HARRIS, J. ARTHUR, and GORTNER, R. A., 1913 On the relationship between the weight of the sugar beet and the composition of its juice. Jour. Ind. Eng. Chem. **5**: 192-195.

1918 The correlation between egg production during the various periods of the year in the domestic fowl. Genetics **3**: 27-72, *figs. 1-11*.

JENKINSON, J. W., 1912 Growth, variability and correlation in young trout. Biometrika **8**: 444-455.

JENNINGS, H. S., 1911 Assortative mating, variability and inheritance of size in the conjugation of Paramecium. Jour. Exp. Zoöl. **11**: 1-134.

LUDWIG, F., 1901 Variationsstatistische Probleme und Materialien. Biometrika **1**: 11-29.

PARKER, G. H., and BULLARD, C., 1913 On the size of litters and the numbers of nipples in swine. Proc. Amer. Acad. Arts and Sci. **49**: 399-426.

PEARL, R., 1909 The frequency constants of a variable $z = f(x_1, x_2)$. Biometrika **6**: 437-438.

STUART, W., 1912 Mendelian inheritance in the carnation. Bull. Vt. Agr. Exp. Sta. 163.

YULE, G. U., 1914 Fluctuations of sampling in Mendelian ratios. Proc. Cambridge Phil. Soc. **17**: 425-432.

Information for Contributors

The length of contributions to GENETICS will not be arbitrarily limited, but articles less than fifty pages may be expected to appear more promptly than those of greater length. The limitations as to subject-matter will be construed broadly, so that any paper of sufficient importance, whether in the field of genetics proper, of cytology, embryology, physiology, biometry, or mathematics, if of such character as to be of primary interest to the geneticist, will find here an appropriate place of publication.

Contributors are requested to observe the usual care in the preparation of manuscripts. All references to the literature should cite the name of the author, followed the year of publication, the papers so referred to being collected into a list of "Literature cited" at the end of the article. In this list great care should be taken to give titles *in full*, and to indicate accurately in Arabic numerals, the volume number, the first and last pages and the date of publication of each paper, if published in a periodical, and the number of pages, place and date of publication and name of publisher, of each independent publication. The arrangement of this list should be alphabetical by authors, and chronological under each author.

Footnotes should be inserted immediately after the reference marks which refer to them, and should be separated from the text by lines. All footnotes should be numbered consecutively in a single series.

No limitation will be placed on the kind of illustrations to be used, but as funds for illustration are still very limited, authors are requested to practice restraint, and to limit themselves to such amount and grade of illustration as seem positively essential to an adequate understanding of their work. The size of the type-forms of GENETICS is $4\frac{5}{8} \times 7\frac{1}{2}$ inches and all illustrations intended for use as single plates should not exceed $5 \times 7\frac{3}{4}$ inches, or should allow of reduction to these dimensions. Double plates may be $12\frac{1}{2} \times 7\frac{3}{4}$ inches. Legends for all figures should be typewritten on a separate sheet.

Authors will receive galley-proofs, and, if specially requested, page-proofs will be sent also, it being understood that the latter must be returned very promptly, and that no extensive change may be made in page-proofs, which is not compensated for within the same paragraph, or in an adjacent paragraph on the same page.

Contributors will receive 100 reprints of their papers, with covers, free of charge, and may secure additional copies at actual cost of manufacture, provided these are ordered when the corrected galley-proofs are returned.

Manuscripts and all editorial correspondence should be addressed to the Managing Editor, Dr. George H. Shull, 60 Jefferson Road, Princeton, N. J.

GENETICS, JULY 1918

TABLE OF CONTENTS

PRACTICAL UNIVERSALITY OF FIELD HETEROGENEITY AS A FACTOR INFLUENCING PLOT YIELDS

BY

J. ARTHUR HARRIS

PRACTICAL UNIVERSALITY OF FIELD HETEROGENEITY AS A FACTOR INFLUENCING PLOT YIELDS

By J. Arthur Harris

Collaborator, Office of Western Irrigation Agriculture, Bureau of Plant Industry, United States Department of Agriculture

INTRODUCTION

With the development of a more intensive agriculture there must be a wider use and a progressive refinement of the method of plot tests in agronomic experimentation. Betterment of the method of plot tests must be sought along two lines, (1) the perfection of biological technic and (2) the more extensive use of the modern higher statistical methods in the analysis of the results.

In 1918 Mr. C. S. Scofield, in charge of the Office of Western Irrigation Agriculture, and Prof. E. C. Chilcott, in charge of the Office of Dry-Land Agriculture, asked the writer to undertake an investigation of the statistical phases of the problem of the accuracy of plot tests. The present paper deals with one aspect only of the general problem, that of the lack of uniformity of the experimental field. This is both the most potent cause of variation in plot yields and the chief difficulty in their interpretation.

Many of the careful writers on field experimentation have noted the existence of soil heterogeneity. Few have, however, sufficiently recognized and none have adequately emphasized the importance of this factor.

The problem of field heterogeneity is twofold. First, some measure of the amount of its influence upon crop yields must be obtained. Second, some means of avoiding or of correcting for its influence must, if possible, be secured.

An exact measure of the influence of field heterogeneity, and not merely a vague notion that it may influence experimental results, is the first and most fundamental step in the closer analysis of the factors determining the variability of plot yields. If the application of such a criterion to results obtained by practised agriculturalists from fields selected for their uniformity shows no evidence of heterogeneity, plot tests may be carried out along conventional lines with confidence that

with reasonable precautions reliable results will be obtained. If, on the other hand, the application of such a criterion shows a high degree of irregularity in fields selected for their uniformity by experienced agriculturalists, it is evident that very special precautions must be taken to obtain trustworthy results. Some quantitative measure, and some probable error of this measure, of the amount of irregularity of the soil of a field, as shown by actual capacity for crop production, and not merely a demonstration of its existence is, therefore, required.

The purpose of this paper is to show by the analysis of the actual yields of test plots reported by agricultural experts that the securing of fields suitable for a direct comparison of yields is, practically speaking, an impossibility. The results show that unless special precautions are taken irregularities in the field may have greater influence upon the numerical results of an experiment than the factors in crop production which the investigator is seeking to compare.

The results of this study may seem to be altogether negative—destructive rather than constructive. The unbiased student must, however, admit that a full evaluation of all the sources of error is an essential prerequisite to constructive work. Furthermore, large expenditures of public funds are being devoted to fertilizer tests, variety tests, and rotation experiments. It is preeminently worth while to ascertain to what extent results derived from methods now in use may be considered reliable.

Subsequent papers will treat other phases of the problem.

FORMULAE

A criterion of field homogeneity (or heterogeneity) to be of the greatest value should be universally applicable, be comparable from species to species, character to character, or experiment to experiment, and be easy to calculate.

In 1915 the suggestion was made (5)[1] that we may proceed as follows: Suppose a field divided into N small plots, all sown to the same variety of plants. Let p be the yield of an individual plot. The variability of p may be due purely and simply to chance, since the individuals of any variety are variable and the size of the plots is small, or it may be due in part to the diversity of conditions of the soil. If irregularities in the experimental field are so large as to influence the yield of areas larger than single plots,[2] they will tend to bring about a similarity of adjoining plots, some groups tending to yield higher than the average, others lower.

Now let the yields of these units be grouped into m larger plots, C_n, each of n contiguous ultimate units, p. The correlation between the

---

[1] Reference is made by number (italic) to "Literature cited," p. 313–314.

[2] Irregularities of soil influencing the plants of only a single small plot may in most work be left out of account, since they are of the kind to which differences between individuals are to a considerable extent due and are common to all the plots of a field.

p's of the same combination plot, C_n, will furnish a measure (on the scale of o to $\pm$ 1) of the heterogeneity of the field as expressed in capacity for crop production. If this correlation be sensibly o (under conditions such that spurious correlation is not introduced), the irregularities of the field are not so great as to influence in the same direction the yields of neighboring small plots. As heterogeneity becomes greater the correlation will also increase. The value of the coefficient obtained will depend somewhat upon the nature of the characters measured, somewhat upon the species grown, somewhat upon the size of the ultimate and combination plots, and to some degree upon the form of the combination plots.

Knowledge of the values of the correlations to be expected must be obtained empirically.

Let S indicate summation for all the ultimate or combination plots of the field under consideration, as may be indicated by C_n or p. Let $\bar{p}$ be the average yield of the ultimate plots and σ_p their variability, and let n be constant throughout the m combination plots. Using the formulae of an earlier memoir (3) in a notation which is as much simplified as possible for the special purposes of this discussion,

$$r_{p_1 p_2} = \frac{\{[S(C_n{}^2) - S(p^2)] / m[n(n-1)]\} - \bar{p}^2}{\sigma_p{}^2}.$$

This formula assumes the combination plots to be of uniform size— that is, to contain each the same number, n, of ultimate plots. It may be desirable or necessary to have some of the combination plots smaller than the others.

Such cases are frequently met in practical work. For example, the wheat field of Mercer and Hall is laid out in a 20 by 25 fold manner. This permits only 2 by 5, 4 by 5, or 5 by 5 combinations of the same size throughout. One of Montgomery's experiments with wheat covered an area of 16 by 14 plots which may be combined in only 2 by 2 or 4 by 2 fold groupings to obtain equal areas suitable for calculation. In each of these cases other groupings are desirable.

The formulae are quite applicable to such cases; the arithmetical routine is merely a little longer. The formula is as above, but $\bar{p}$ and σ_p are obtained by a $(n-1)$-fold weighting of the plots,[1] where n is the variable number of ultimate plots in the combination plot to which any p may be assigned—that is,

$$\bar{p} = S[(n-1)p] / S[n(n-1)],$$

$$\sigma_p{}^2 = \frac{S[(n-1)p^2]}{S[n(n-1)]} - \left(\frac{S[(n-1)p]}{S[n(n-1)]}\right)^2.$$

[1] That is, each ultimate plot is multiplied by the number less one of the plots in the combination plot to which it is assigned.

Ample illustration of the arithmetical routine has been given in the original paper.

The formulae employed assume the symmetry of the correlation surface. It has been shown elsewhere (4) that spurious values of the correlation coefficient may arise in such cases. Since both $\bar{p}_1\bar{p}_2$ and $\sigma_{p_1}\sigma_{p_2}$ take the maximum values when, because of the symmetry of the correlation surfaces, $\bar{p}_1 = \bar{p}_2$, $\sigma_1 = \sigma_2$, it is clear that the limiting value of the spurious correlation will be 0.

Thus it is possible that heterogeneity exists even when $r_{p_1p_2} = 0$, but a field can not be considered homogeneous if $r_{p_1p_2}$ has a value which is statistically significant in comparison with its probable error.

Practically, little difficulty will arise from this source, and it can usually be easily avoided by the exercise of a little care in the selection of the proper grouping in doubtful cases.

According to the foregoing conception the relationship between the yield of associated plots is expressed on the universally comparable scale of r, ranging from 0 to ± 1.

When symmetrical tables are used—that is, when each plot is used once as a first and once as a second member of the associated pair—$\bar{p}_1 = \bar{p}_2$, $\sigma_{p_1} = \sigma_{p_2}$, and the regression slope is identical with the correlation coefficient.

Thus, if one ultimate plot, p_1, of a combination plot be known, the most probable deviation of another plot will be $p_2 - \bar{p} = (p_1 - \bar{p})r$.

Concretely, if the yield of a first plot of a combination plot be 10 pounds above the average of the field as a whole and if the interplot correlation be $r_{p_1p_2} = 0.60$, the most probable yield of a second plot will be 6 pounds above the average.

Similar reasoning applies throughout. Those who have difficulty in thinking in terms of correlation coefficients can most easily grasp the significance of the results by remembering that in this case the correlation coefficients multiplied by 100 gives the most probable percentage of deviation of the yield of an associated plot when the deviation of one plot of the group from the general average is known.

INFLUENCE OF SOIL HETEROGENEITY ON YIELD OF FIELD CROPS

In the paper in which these formulae were suggested it was shown that yield of straw and grain and the nitrogen content of wheat, yield of roots and tops of mangolds, and yield of timothy hay are markedly influenced by irregularities in the carefully selected fields upon which plot cultures have been carried out by agriculturalists.

We have now to ascertain whether this is a general phenomenon or whether it is merely a chance result of these particular cultures. The suggestion has been made that the latter is the case, that with the exercise of a little care uniform fields may be secured, and that substratum

heterogeneity was overemphasized as a factor influencing plot tests. This question can be answered only by actually determining the degree of heterogeneity existing in the fields which have passed the criticism of agricultural experts.

It will be conducive to brevity to have a definite system by which the arrangement of the plots in a field may be described. We shall consider the plots arranged as soldiers in ranks and files. The worker inspects the plot records of a field as recorded on a map or table. By ranks we understand the horizontal rows of plots, by files the vertical rows.

1.80	1.83	2.00	1.91	1.90	1.89	1.79	1.75	2.03	1.83	2.18	1.93	1.77	1.86
1.80	2.07	1.77	1.90	1.70	1.79	1.90	2.04	1.95	1.83	2.06	1.76	1.86	1.79
1.93	1.96	1.83	1.92	1.69	1.90	1.80	1.89	1.83	1.85	2.00	2.13	1.82	1.83
1.89	1.96	1.92	1.86	1.79	1.86	1.79	1.94	1.92	1.80	1.97	2.00	1.87	1.73
2.00	2.01	1.89	1.77	1.97	1.85	1.97	2.10	1.99	1.83	2.00	1.92	1.79	1.89
1.96	1.96	2.00	1.82	1.93	1.82	1.87	1.87	1.92	1.99	1.87	1.83	1.92	1.96
1.89	2.11	1.99	1.87	1.86	1.84	2.06	1.90	1.90	1.82	1.81	1.97	1.79	1.89
2.03	1.86	1.80	1.86	2.06	1.72	1.86	1.72	2.07	1.82	1.84	1.97	1.96	2.01
1.83	1.82	1.82	1.75	1.77	1.72	1.90	1.83	1.90	1.83	1.90	1.85	1.76	2.07
1.87	2.14	1.96	1.87	1.97	1.90	1.90	2.13	1.80	1.83	1.90	2.06	1.94	1.87
1.90	1.94	1.94	1.77	1.89	1.86	1.82	1.87	1.80	1.84	1.87	2.04	1.94	1.89
1.94	1.76	1.96	1.99	1.87	2.04	1.93	1.77	1.74	1.89	1.93	1.96	2.04	1.97
1.83	1.99	1.97	2.08	1.99	1.96	2.15	1.82	1.78	1.83	1.98	1.89	1.85	1.87
1.85	1.87	1.85	1.82	1.92	1.89	2.13	1.82	1.73	1.83	1.96	2.04	1.86	2.08
2.10	1.83	1.85	1.96	2.01	1.92	1.68	1.89	1.85	1.85	1.83	1.85	2.07	1.75
1.93	1.86	1.93	1.87	1.90	1.86	1.99	1.89	1.83	1.82	1.96	1.99	1.99	2.06

Fig. 1.—Montgomery's diagram of 5.5 by 5.5 foot plots of Turkey wheat, showing variations in the percentage of nitrogen in the grain.

Thus figure 1, showing the nitrogen content of wheat plots 5.5 by 5.5 feet given by Montgomery (17), may be considered made up of 16 ranks and 14 files.

In considering rearrangements or combinations of plots we shall refer to the ranks and then to the files—an order easily carried in mind by remembering the trite expression "rank and file." Thus in referring to a 2 by 5 fold combination we mean that two adjacent ranks and five adjacent files of plots were combined. Individual plots may be easily designated. Thus, the plot belonging to the sixth rank[1] and the fifth file in the nitrogen contents of wheat yields contained 1.93 per cent nitrogen.

[1] Ranks are numbered from the top of map, files from the left.

1.—MANGOLDS

The yields of 200 plots of mangolds studied by Mercer and Hall (*15*) may be grouped into combination plots in a 2 by 2 fold manner. When this is done, the correlation between the yields of associated plots has been shown[1] to be as follows:

For weight of roots, $r = 0.346 \pm 0.042$,[2] $r/E_r = 8.24$.
For weight of leaves, $r = .466 \pm .037$, $r/E_r = 12.5$.

Thus, if one plot of a combination plot is higher or lower than the general average by a given amount, an associated plot may be expected to deviate from the general average by 35 to 40 per cent of this amount.

2.—POTATOES

Lyon (*14*) gives the yield in pounds for each of six sections of a series of 34 rows of potatoes. This crop was harvested from "a piece of apparently uniform land." Each section was 72 feet 7 inches in length. The distance between rows was 34 inches.

Combining yields of rows and of sections of rows by twos, we reduce the field from a 34 by 6 fold to a 17 by 3 fold combination. The correlations between the sections of the rows is then found to be

$$r_{p_1 p_2} = 0.311 \pm 0.043, \ r/E_r = 7.30.$$

Yield of potatoes in this field is, therefore, markedly influenced by irregularities of soil conditions.

For data on a second test on the influence of field heterogeneity on the yield of potatoes we may avail ourselves of the valuable records of yields of individual hills reported by Stewart (*19*). Since these are recorded in quadruplets for the purpose of determining the influence of missing hills upon yield,[3] it is not feasible to group them into plots. The influence of heterogeneity may be tested by determining the correlation between the yields of the plants of a quadruplet.[4]

---

[1] For original data see Mercer and Hall (*15, p. 109*); also Harris (*5, p. 434–436*).

[2] The probable errors have in all cases been computed on the basis of the actual, not the weighted, number of ultimate plots as N.

[3] The planting scheme adopted was

     0   a_1 a'_1 b'_1 b_1 0 a_2 a'_2 b'_2 b_2 0 a_3 a'_3 b'_3 b_3 . . . ,

where a and a' are the two halves of the same tuber and b and b' are two halves of another tuber. Thus halves a and b were grown adjoining missing hills and were subject to competition on one side only, whereas halves a' and b' were subject to competition from two adjacent plants.

[4] Since a and a' are halves of the same tuber and b and b' are halves of another, the correlations $r_{aa'}$, $r_{bb'}$ might be due to a specific physiological influence of the characters of the tuber upon both plants developing from the corresponding half tubers rather than to an influence of differences in soil conditions. We have, therefore, determined the correlations between the plants occupying the same relative position in the quadruplet but derived from different parent tubers, that is r_{ab}, $r_{a'b'}$. Hence r_{ab} represents the correlation between the two outside tubers and $r_{a'b'}$ the correlation between the two inside tubers of the quadruplet. As a control on the results these correlations between one outside and one inside plant have been determined. These are $r_{ab'}$ and $r_{ba'}$.

The data given by Stewart are number of tubers and total weight of tubers per plant. These two characters permit the determinations of the average weight per tuber.

When all the pairs are omitted which have been omitted by Stewart[1] or have been designated as affected by leafroll, there remain 139 quadruplets. Determining the correlations between the yield of the two plants derived from different tubers but exposed to the same conditions for growth, we have the following correlations:

For number of tubers per hill—

$$r_{ab} \ \ = 0.318 \pm 0.051, \ r/E_r = \ 6.19.$$
$$r_{ab}' \ = \ .138 \pm \ .056, \ r/E_r = \ 2.46.$$
$$r_{a'b} \ = \ .230 \pm \ .054, \ r/E_r = \ 4.26.$$
$$r_{a'b}' \ = \ .220 \pm \ .054, \ r/E_r = \ 4.04.$$

For total weight of tubers per hill—

$$r_{ab} \ \ = 0.457 \pm 0.045, \ r/E_r = 10.10.$$
$$r_{ab}' \ = \ .312 \pm \ .052, \ r/E_r = \ 6.00.$$
$$r_{a'b} \ = \ .427 \pm \ .047, \ r/E_r = \ 9.09.$$
$$r_{a'b}' \ = \ .290 \pm \ .052, \ r/E_r = \ 5.53.$$

For average weight of tubers—

$$r_{ab} \ \ = 0.237 \pm 0.054, \ r/E_r = \ 4.39.$$
$$r_{ab}' \ = \ .104 \pm \ .057, \ r/E_r = \ 1.82.$$
$$r_{a'b} \ = \ .054 \pm \ .057, \ r/E_r = \ \ .95.$$
$$r_{a'b}' \ = \ .117 \pm \ .056, \ r/E_r = \ 2.07.$$

The correlations are positive throughout and generally statistically significant with regard to their probable errors. They show, therefore, that this experimental plot was heterogeneous to an extent that influenced in a very measurable degree the number of tubers, the total weight of tubers, and the average weight of tubers of neighboring hills. For all four measures of interdependence the coefficients are lowest for average weight of tubers and highest for total weight of tubers, while the correlations for number of tubers produced are intermediate in value.

The values of r_{ab} are consistently higher than those for $r_{a'b'}$, notwithstanding the fact that a' and b' are more closely associated than a and b. The measures of interrelationship between the yields of pairs of plants, one of which occupies an inside and the other an outside position in the quadruplet, are sometimes intermediate between r_{ab} and $r_{a'b'}$ and sometimes less than $r_{a'b'}$. On the assumption that the correlation is due solely to environmental influence one would expect the highest

[1] Records have been abstracted from Stewart's Table I. Prof. Stewart has kindly furnished some additional information in regard to certain entries in this table.

correlation between the most closely associated plants—that is $r_{a'b'} > r_{ab}$. Apparently the reverse condition, $r_{a'b'} < r_{ab}$, is due to some influence of the open space adjoining a and b, which allows the fuller development of those plants and in consequence renders them more representative of the extremely localized soil influences to which they are subjected.[1]

III		II	
b	a	b	a
230	305	290	305
180	290	240	290
200	310	300	340
210	265	285	355
200	260	300	325
225	285	280	345
215	285	275	365
220	235	270	285
255	235	285	285
210	230	280	260
240	245	300	285
235	235	265	265
230	270	270	295
210	260	270	285
225	260	315	340
225	235	320	330
220	240	275	315
230	200	285	350
255	225	295	340
265	255	310	295
235	225	320	305
250	280	310	315
240	265	310	280

FIG. 2.—Diagram showing yield of alfalfa in first cutting, 1913, on the Huntley experimental tract. The yield is expressed in pounds per half plot.

3.—TIMOTHY HAY

The records of plot yields of timothy hay published by Holtsmark and Larsen (*8*) have been shown elsewhere (*5*) to present a correlation between the yield of ultimate plots, combined in a 2 by 2 fold manner, of

$$r = 0.611 + 0.027, r/E_r = 22.4.$$

Clearly the field was highly heterogeneous.

4.—ALFALFA HAY

Records of the yields of a series of 46 plots on the Huntley Experiment Farm, Montana, may be used to test further the influence of heterogeneity on the yields of alfalfa hay. Data were kindly placed at my disposal by Mr. C. S. Scofield.

Alfalfa should be of especial interest in the present discussion since it is a deep-rooted perennial herb, whereas all other herbaceous crops investigated have been annuals, or at most biennials.

In field B of this experimental farm there are two series, II and III, each of 23 plots. The 46 plots form a solid block which has been planted each year to one crop just as if it were an ordinary field. The two series of plots are separated from each other only by a temporary irrigation ditch. Each plot is $23\frac{1}{3}$ feet wide, 317 feet long, and contains approximately 0.17 acre. These plots have in certain cases been harvested in subplots of 0.085 acre when the division has been into halves, of 0.0567 acre when the division has been

---

[1] Possibly competition between closely associated a' and b' plants tends to make the yield of one low when that of the other is high.

into thirds, and of 0.0425 acre when the division has been into quarters of plots.

In the spring of 1912 the whole field was uniformly seeded to alfalfa; only one crop was harvested, and yields were recorded for the entire

III				II			
b		a		b		a	
70	95	125	135	135	155	135	175
110	75	85	160	145	125	125	165
80	90	125	110	165	155	150	160
100	65	130	130	145	180	145	180
115	95	110	125	135	165	100	140
115	125	135	135	125	185	130	155
110	95	120	115	145	175	100	155
120	90	100	115	140	150	100	180
100	90	80	105	125	150	45	150
95	95	105	120	125	140	60	145
115	80	95	100	120	140	65	110
115	90	90	105	125	145	120	60
110	100	110	130	120	140	110	115
115	85	120	165	130	150	100	130
105	105	100	145	130	150	145	140
150	95	100	95	100	150	110	115
135	115	90	105	95	110	100	130
155	125	120	100	65	130	115	115
145	130	145	95	120	120	100	115
170	135	155	105	95	135	95	115
135	125	155	95	110	120	115	110
140	115	160	120	110	145	115	130
150	100	120	160	85	150	105	85

Fig. 3.—Diagram showing yield of alfalfa in second cutting, 1913, on the Huntley experimental tract. The yield is expressed in pounds per quarter plot.

plots only. In 1913 and 1914 three cuttings were made. The first cutting was harvested in half plots. The second cutting of 1913 and the first and second cuttings of 1914 were harvested in quarter plots. The

third cutting of 1913 was lost because of a heavy wind which mixed the plot yields at harvest time, so that it was implossible to secure

III				II			
b		a		b		a	
85	85	130	120	130	150	140	165
105	100	105	120	135	150	140	185
100	80	105	110	120	150	170	165
105	110	95	130	165	155	150	170
100	100	105	130	120	140	145	185
100	105	100	125	120	175	195	155
90	100	100	120	155	155	115	200
90	100	105	120	85	155	145	170
120	95	90	120	115	140	170	165
85	95	75	110	155	130	105	155
75	95	85	105	85	130	125	240
60	110	90	100	120	140	160	135
75	100	75	140	95	120	120	130
55	100	75	140	120	130	125	165
75	95	85	125	120	130	140	145
85	100	60	115	125	120	140	160
85	105	100	105	120	135	135	150
115	100	65	115	115	140	155	130
115	125	85	125	150	125	140	130
85	135	95	120	135	135	135	135
105	120	105	105	130	140	165	145
100	115	125	135	140	160	170	140
100	115	140	120	135	120	115	120

FIG. 4.—Diagram showing yield of alfalfa in first cutting, 1914, on the Huntley experimental tract. The yield is expressed in pounds per quarter plot.

accurate weights on any of the plots. The third cutting for 1914 was harvested in subplots one-third the size of the original plots.

The actual yield of these subdivisions is indicated in figure 2[1] for the first cutting and figure 3 for the second cutting in 1913 and in figure 4

---

[1] Diagrams are set in type instead of being drawn to scale.

for the first cutting, figure 5 for the second cutting, and figure 6 for the third cutting in 1914.

III				II			
b		a		b		a	
100	110	135	125	120	145	145	140
80	85	110	120	130	145	175	155
70	110	140	115	170	155	195	170
70	140	115	125	160	190	145	165
85	125	85	125	180	190	155	175
55	125	95	100	190	175	185	185
65	105	115	115	225	155	200	195
65	110	95	110	190	190	180	165
70	105	100	135	140	155	155	165
110	120	60	100	110	120	100	175
100	110	85	125	95	125	70	140
95	120	120	95	75	100	145	105
110	135	125	135	100	75	125	145
130	120	95	150	135	85	90	170
115	115	100	140	115	125	105	170
130	130	80	115	95	110	95	140
135	115	65	110	110	85	90	150
110	115	80	120	120	130	95	180
145	160	75	135	120	125	105	140
140	135	80	125	105	145	155	100
135	135	90	120	115	155	140	125
120	155	110	130	130	130	135	130
90	160	110	115	120	130	120	75

FIG. 5.—Diagram showing yield of alfalfa in second cutting, 1914, on the Huntley experimental tract. The yield is expressed in pounds per quarter plot.

For the yield of alfalfa on quarter plots for the second cutting in 1913 and the first and second cuttings for 1914 and in third plots for the third cutting for 1914 the correlations are

1913, second cutting, $r = 0.182 \pm 0.048$, $r/E_r = 3.79$.

1914, first cutting, $r = 0.432 \pm 0.040$, $r/E_r = 10.7$.
1914, second cutting, $r = .449 \pm .040$, $r/E_r = 11.3$.
1914, third cutting, $r = .311 \pm .052$, $r/E_r = 5.99$.

III			II		
x	y	z	x	y	z
230	190	225	160	240	180
220	170	130	220	220	165
215	150	130	200	205	190
175	150	115	205	190	215
175	155	125	205	220	170
155	155	105	175	160	175
190	130	125	160	175	165
155	145	115	170	165	165
170	105	110	160	155	160
140	120	100	150	120	180
155	90	140	95	160	145
125	125	120	125	165	155
210	100	125	145	160	150
175	140	110	180	165	140
155	145	155	180	195	165
140	115	155	165	185	125
150	125	155	170	170	120
115	120	150	170	150	135
160	150	165	150	165	150
140	165	140	150	165	160
155	155	155	165	195	150
150	175	170	175	160	185
185	150	140	90	155	135

FIG. 6.—Diagram showing yield of alfalfa in third cutting, 1914, on the Huntley experimental tract. The yield is expressed in pounds per third plot.

It will be noted that the results are in very close agreement indeed for 1914. The second cutting for 1913 differs significantly from the others, but no explanation can be suggested.

Grouping all yields in two comparable subplots, we find

1913, first cutting, $r = 0.407 \pm 0.059$, $r/E_r = 6.93$.
1913, second cutting, $r = .343 \pm .062$, $r/E_r = 5.52$.
1914, first cutting, $r = .602 \pm .045$, $r/E_r = 13.4$.
1914, second cutting, $r = .657 \pm .040$, $r/E_r = 16.4$.

We note that all the correlations are higher for a 2-fold division than for a 4-fold division. The coefficients for the second cutting of 1913 are again lower than the other values.

The foregoing results are based upon weightings of single cuttings only. It is now desirable to determine the correlations for yield of first and second cuttings combined.

If the combined yield be considered in quarter plots as ultimate units in 1914 we find

$$r = 0.517 \pm 0.036, \ r/E_r = 14.2.$$

Combining to obtain total yield in half plots in both 1913 and 1914, we have the following correlations between the yields of the two half plots:

For 1913, $r = 0.387 \pm 0.060$, $r/E_r = 6.46$.
For 1914, $r = .709 \pm .035$, $r/E_r = 20.2$.

5.—STRAW AND GRAIN IN WHEAT

The data of the Rothamsted wheat plots,[1] analyzed in an earlier paper (*5, p. 436–440, 443–444*), show the following correlations when the 500 plots are grouped in 2 by 2 fold manner for the first 22 files and in a 2 by 3 fold manner for the twenty-third to the twenty-fifth file:

For yield of grain, $r = 0.336 \pm 0.027$, $r/E_r = 12.5$.
For yield of straw, $r = .483 \pm .023$, $r/E_r = 20.9$.

6.—STRAW AND GRAIN IN RAGI, ELEUSINE CORACANA

Lehmann (*12*) has given a series of data derived from the yields of grain and straw of ragi cultivated on the dry-land tract of the Experimental Farm at Hebbel, near Bangalore, Mysore State. The plots used were of 1/10-acre area.

The land was previously owned by several raiyats who have naturally treated it somewhat differently in regard to manuring and cultivation. The various pieces used as garden lands are of course in much better condition than those used for ordinary dry crops. This causes considerable temporary differences to exist in some of the plots in addition to probably slight permanent differences. (*12, 6th Rpt., p. 2.*)

From these conditions one would expect a high degree of heterogeneity in the series of plots. The data permit the testing of the possibility of a decrease in heterogeneity due to uniformity of crop and treatment for three years.

---

[1] For data see Mercer and Hall (*15, p. 119*); also Map B of Harris (*5*).

These data are, furthermore, of particular interest since they consist of the records of yields for three successive years of the same crop on a series of unirrigated plots in a region where crop production is subject to many uncertainties because of inadequate rainfall.

Fortunately for our present purposes the meteorological conditions during the three years covered by this experiment were very different from year to year. The values of the most significant factor, the July to October rainfall, are given in Table I. This shows that the rainfall in 1906 was practically twice as heavy as in either of the other two years.[1]

TABLE I.—*Rainfall at Hebbel, near Bangalore, Mysore State, India*

Month.	1905	1906	1907	Average of 10 years.
	Inches.	Inches.	Inches.	Inches.
July...	1. 77	7. 09	4. 17	3. 04
August.......................................	6. 75	9. 98	1. 50	4. 32
September...................................	1. 47	5. 50	5. 66	8. 14
October......................................	5. 76	8. 51	.81	5. 97
Total.................................	15. 75	31. 08	12. 14	21. 47

Maps of the fields are given in the sixth annual report for 1904–1905. Further descriptive detail is given in the seventh, eighth, and ninth reports for 1905–1908. The yield of grain and straw in plots of 1/10 acre grown in 1905 is given in the seventh report. The eighth report gives detail of the crop of 1906 but does not contain the yields, which are summarized for the years 1905, 1906, and 1907 in Tables I and II of the ninth report.

Unfortunately the yields of a considerable number of the plots have had to be omitted from maps I and II of Lehmann's report. In combining in a 2 by 2 fold manner it is necessary either to disregard all combination plots in which there are not four ultimate plots or to weight properly in using those containing 2 or 3 plots only. The course followed has been to group the plots by fours and to determine the correlation by the formulae for a variable number of plots when all of the ultimate plots were not planted.

The following table shows the correlation between the yield of grain, of straw, and of grain and straw:

	1905	1906	1907
Grain.......................................	0. 735±0. 031	0. 138±0. 065	0. 716±0. 032
Straw.......................................	.424± .055	.164± .065	.573± .045
Total yield.................................	.415± .055	.145± .065	.636± .040

[1] A discussion of the growth of these crops in relation to the distribution of the rainfall appears in Lehmann's ninth report (*12, p. 2–7*).

The results are of unusual interest.   In 1905 and 1907 the correlation between yields of grain are unusually high, falling only slightly below three-fourths of perfect correlation.   The correlations for yields of straw and for both grain and straw are of medium value in those two years. In 1906, however, the correlations for all the characters are of a very low order; and any one of them taken alone might not be considered significant in comparison with its probable error, which has been calculated on the basis of 103 plots, the number actually involved in the calculations.

Apparently the unusual moisture conditions of 1906 tended to obliterate the differences in the field to which the individuality of adjoining plots was due.

That the unusual weather had a profound influence on the yield of the plots is shown by Table II, in which the means, standard deviations, and coefficients of variation for the yield of the individual plots are set forth.[1]

TABLE II.—*Means, standard deviations, and coefficients of variation for the yield of ragi at Hebbel, near Bangalore, Mysore State, India*

[Yield expressed in pounds per 1/10 -acre plot]

Year.	Grain.			Straw.			Total yield.		
	Mean.	Standard deviation.	Coefficient of variation.	Mean.	Standard deviation.	Coefficient of variation.	Mean.	Standard deviation.	Coefficient of variation.
1905..........	192. 8	31. 5	16. 3	360. 8	148. 8	41. 2	553. 5	190. 3	34. 4
1906..........	136. 6	47. 1	34. 5	191. 6	82. 0	42. 8	328. 1	127. 4	38. 8
1907.........	165. 0	48. 3	29. 3	295. 4	80. 2	27. 1	460. 4	126. 9	27. 6

The means show that yield of both grain and straw was much lower in the abnormally wet year than in either of the others.   The standard deviations are of course largely influenced by the actual magnitudes of the yields and are, in consequence, difficult of interpretation.   The relative variabilities, as measured by the coefficients of variation, are more orderly.   They show that for grain, straw, and total yield the variability of the individual plot yields is greater in the wet year.

Thus the influence of the wet season has not been to make the yield of all the plots alike.   It has tended to decrease yield and to increase relative variability from plot to plot.   But at the same time it has tended to screen certain factors which in drier years have a marked influence on the individuality of the plots.

Further analysis is not desirable without more detailed information concerning the plots.   From the information at hand it seems quite

---

[1] These constants are obtained by weighting in an (n–1)-fold manner, since this was the method followed in obtaining the constants for the heterogeneity coefficient.

clear that the innate differences in different parts of the field do not in some seasons exert their full influence upon crop yield because of the weight of other factors.   The practical conclusion to be drawn from this result is that an experimental field which might be demonstrated to be sensibly uniform for one crop plant or for one season might not prove to be so for another crop or in a different season.

7.—KHERSON OATS

Kiesselbach (*10, 11*) has given records of yield for 207 1/30-acre plots of Kherson oats.   He says:

These plats were planted   .   .   .   upon a seemingly uniform field for the purpose of studying variation in plat yield as a source of experimental error.   The entire field had been cropped uniformly to silage corn for a period of eight years.   It had been plowed each year and was also plowed in preparation for the oats in 1916.   The oats were drilled during two successive days in plats 16 rods by 66 inches   .   .   . The plats were separated by a space of 16 inches between outside drill rows.   A wide discard border of oats was grown around the outer edge of the field, so that all plats should have a similar exposure.

Love (*13*) has shown the existence of heterogeneity in this field.

Grouping the entries of Kiesselbach's Table 27 in a 3 by 1 fold manner the heterogeneity coefficient is found to be

$$r = 0.495 \pm 0.035, \; r/E_r = 14.$$

For data on a second test of the influence of heterogeneity on the yields of experimental plantings of oats we turn to a small experiment by Montgomery (*17*), who has given the yields of thrashed grain in grams from 100 consecutive rows of Kherson oats (*17, p. 35, Table XIII*) each 12.5 feet in length.

The plat chosen for this test was quite uniform and the appearance of the plat at harvest was very satisfactory.

Combining by twos, we find for the correlation between adjacent rows

$$r = 0.339 \pm 0.060, \; r/E_r = 5.65.$$

8.—GRAIN AND NITROGEN CONTENT IN WHEAT

Montgomery (*17, p. 37, fig. 10*) has given the yield of grain in grams on 224 blocks each 5.5 feet square.   Combining in a 2 by 2 fold manner we deduce

$$r = 0.391 \pm 0.038, \; r/E_r = 10.2.$$

Again, Montgomery (*17, p. 21-22, fig. 7*) has given the values of nitrogen content from 224 Turkey wheat plots of the same size.   These values are quoted in figure 1 of this paper.   The correlation between the plots is found to be

$$r = 0.020 \pm 0.045, \; r/E_r = 0.44.$$

Finally, Montgomery (*16*) has given data for both yield of grain and nitrogen content on 224 plots of wheat grown at the University of Nebraska in 1911. The plot (77 by 88 feet) had been sown continuously to Turkey winter wheat for three years.

The plat was of about average uniformity and fertility.

When grouped in a 2 by 2 fold manner these plots of wheat have been shown (*5, p. 440–441, map C*) to give the following correlations:

For yield of grain, $r = 0.603 \pm 0.029$, $r/E_r = 21$.
For percentage of nitrogen, $r = .115 \pm .044$, $r/E_r = 2.59$.

Yield of grain per plot is clearly influenced by irregularities of the experimental field, notwithstanding the fact that the plots are only 5.5 by 5.5 feet in area. The correlation for percentage of nitrogen is not certainly significant.

9.—HOPS

Stockberger (*20*) has given a series of yields for 30 rows of hops which he believes to be quite typical of many thousands of acres in the Sacramento Valley in California. The yields of these rows cover the period of 1909 to 1914. Combining the rows by twos and determining the correlation between the yield of the adjacent rows of the 15 pairs for each of the years, we obtain the following constants:

Year.	Correlation.	r/E_r.
1909	0.444 ± 0.099	4.50
1910	.695 ± .064	10.91
1911	.061 ± .123	8.50
1912	.326 ± .110	2.97
1913	.606 ± .078	7.79
1914	.386 ± .105	3.69
Average	.419	5.06

Without exception the coefficients are positive in sign. In general they are fairly large and indicate a substantial degree of heterogeneity in this limited area. Probably the heterogeneity would have been shown to be greater had it been possible to work with yields from the sections of the long rows instead of with the rows as a whole.

10.—UNHUSKED RICE

Coombs and Grantham (*2*) give the yield in gantangs of a series of 54 square plots ½ by ½ chain in dimension.

These plots are arranged in 18 ranks and 3 files. They were harvested from a field of standing rice on which—

the crop was extremely regular, as judged before the cutting, and it had not been subjected to any attack of borer or any devastation of rats or birds.

The yields of the original plots are shown in figure 7. These may be combined in a 2 by 1 fold manner to give a correlation of

$$r = 0.344 \pm 0.081, \; r/E_r = 4.25.$$

These rice yields taken from a field described as "extremely regular" show that as a matter of fact the field is heterogeneous and that this **irregularity** influences in a measurable degree the yields of the plots.

13. 6	12. 0	11. 4
14. 6	14. 0	12. 2
14. 8	14. 4	12. 0
13. 0	12. 4	12. 8
15. 0	12. 0	12. 0
13. 4	13. 8	14. 0
14. 2	12. 2	13. 0
14. 0	12. 0	12. 8
14. 0	12. 0	13. 4
14. 0	14. 0	12. 4
15. 0	14. 0	12. 6
14. 8	14. 0	12. 4
14. 0	14. 0	12. 0
14. 4	13. 6	12. 4
12. 6	13. 0	12. 0
12. 2	14. 0	12. 8
11. 6	12. 0	11. 8
12. 4	14. 0	12. 4

FIG. 7.—Diagram showing yield of unhusked rice on Coombs and Grantham's 54 plots ⅓ by ⅓ chain square. The yield is expressed in gantangs per plot.

II.—EAR CORN

Smith (*18*) has published a series of corn yields for three years on plots of $^{1}/_{10}$ acre. The yields are given in his original paper. He has kindly supplied the map showing the relative positions of these plots, which are arranged thus:

101, 201, ..., 601
102, 202, ..., 602
· , · , , ·
· , · , , ·
· , · , , ·
120, 220, ..., 620

Combining yields in a 2 by 1 fold manner, we find for the correlation between the yields of adjacent $^1/_{10}$-acre plots

For 1895, $r = +0.830 \pm 0.019$, $r/E_r = 43.4$.
For 1896, $r = + .815 \pm .021$, $r/E_r = 39.6$.
For 1897, $r = + .606 \pm .039$ $r/E_r = 15.5$.

It is evident that the field was rather highly heterogeneous.

III				II			
b		a		b		a	
133	132	138	142	136	132	148	140
141	141	132	138	145	135	162	156
135	109	125	135	133	116	147	130
132	153	131	131	130	123	155	150
132	137	135	140	137	112	131	129
135	132	135	131	134	126	126	135
131	128	121	125	126	115	122	136
135	125	128	131	121	115	129	137
133	125	125	130	131	124	129	131
137	124	117	131	127	125	129	132
130	117	119	127	132	129	122	141
134	122	115	125	133	123	119	132
129	122	120	132	130	125	136	137
123	118	125	130	124	124	123	136
129	126	134	129	122	126	127	136
134	124	120	121	126	130	132	136
128	125	115	115	122	123	140	135
128	121	110	110	116	115	125	123
127	124	119	107	114	116	110	115
134	112	121	123	122	126	116	125
145	148	133	125	132	127	126	134
149	154	165	160	162	144	137	130
168	169	165	152	158	169	143	108

Fig. 8.—Diagram showing yield of ear corn, 1915, on the Huntley experimental tract. The yield is expressed in pounds per quarter plot.

For a second test of the influence of field heterogeneity on the yield of ear corn we turn to the Huntley data.

III				II			
b		a		b		a	
78	94	104	128	110	121	132	150
73	81	104	118	116	116	140	142
66	77	84	110	113	102	128	138
66	73	80	99	115	113	128	139
77	79	79	103	116	118	126	127
71	73	86	82	100	110	108	132
76	59	86	90	110	117	111	151
94	65	86	100	102	105	118	116
98	75	80	100	111	101	104	118
88	76	74	99	108	92	102	113
91	82	69	80	100	97	101	101
97	87	83	90	103	92	88	96
75	81	80	107	96	78	96	106
67	76	73	117	95	70	90	117
98	85	74	103	98	84	100	116
111	88	76	97	97	92	110	110
108	88	73	84	84	86	104	115
115	97	66	89	100	87	98	123
104	120	86	100	94	94	97	119
110	106	92	99	96	100	83	104
118	110	100	98	114	108	113	120
108	100	105	110	93	99	117	104
108	98	95	100	103	99	114	98

FIG. 9.—Diagram showing yield of ear corn, 1916, on the Huntley experimental tract. The yield is expressed in pounds per quarter plot.

In 1915 and 1916 corn was grown on the Huntley experimental plots, described above, and was harvested in quarter plots. The yields for the two series are shown in figure 8 for 1915 and in figure 9 for 1916. These records are of special interest in view of the fact that these are irrigated

fields, whereas the data provided by Smith are based on corn grown without irrigation.

Retaining the original division into quarter plots, we deduce for the correlation between the subplots

$$\text{For } 1915, r = 0.498 \pm 0.037, r/E_r = 13.4.$$
$$\text{For } 1916, r = .436 \pm .040, r/E_r = 10.8.$$

The results for the two years can not, with due regard to their probable errors, be considered to differ significantly. They indicate a degree of heterogeneity in these Huntley plots quite comparable with that of fields planted to various crops by other observers.

If the quarter plots be combined by adjacent twos and the correlation between the half plots be determined, we find

$$\text{For } 1915, r = 0.494 \pm 0.053, r/E_r = 9.29.$$
$$\text{For } 1916, r = .0431 \pm .057, r/E_r = 7.53.$$

The measure of heterogeneity has been only slightly lowered by dividing the plots into halves instead of into quarters.

INFLUENCE OF SUBSTRATUM HETEROGENEITY ON YIELD OF OR-
CHARD CROPS

In the preceding illustrations the crops considered have been herbaceous plants which are generally fairly superficial in their relation to the soil and most of which complete their development in one or two seasons. It seems of particular interest to extend the studies, as Batchelor and Reed (1) have done, to the yield of large individual plants, such as orchard trees.

For the purpose we employ the splendid series of data of Batchelor and Reed. They say of their various groves (1, p. 251):

The fruit plantations herein discussed, to judge by the surface soil, size, and condition of the trees, as well as their apparent fruitfulness, appeal to the observer as uncommonly uniform. All the orchards studied are situated in semiarid regions and are artificially irrigated during the summer months. This fact is believed to be a distinct advantage for the purpose of reducing the variability of one year's yield compared with another, since it insures a fairly uniform water supply for the soil and reduces one of the variants inevitable in nonirrigated localities.

In the case of the Arlington navel oranges grouped in 8–tree plots as the ultimate unit the authors (1, p. 264) report a correlation between plots of $r = 0.533 \pm 0.085$ when the plots are combined by fours.

It has seemed desirable to test the homogeneity of the soil in each of the orchards studied by them. In determining the following coefficients the individual tree has in each case been the ultimate unit.[1]

Consider first the relationship between the yields of adjacent trees of two navel orange groves.

---

[1] Yields are reported in pounds per tree of ungraded product.

Grouping the yield of the 1,000 trees at Arlington, shown in figure 1 of Batchelor and Reed, in a 2 by 2 fold manner we find

$$r = 0.517 \pm 0.016, \; r/E_r = 33.1.$$

A navel orange grove of 495 trees at Antelope Heights, mapped as figure 2 by Batchelor and Reed, when combined in a 3 by 3 fold manner gives

$$r = 0.375 \pm 0.026, \; r/E_r = 14.4.$$

Grouping the 240 Valencia orange trees of the grove shown in figure 3 of Batchelor and Reed in a 2 by 2 fold manner, we find for the correlation between yields

$$r = 0.306 \pm 0.039, \; r/E_r = 7.75.$$

For the yield in pounds per tree of Eureka lemons as shown in figure 4 of the authors cited, we find for a 2 by 2 fold grouping

$$r = 0.448 \pm 0.028, \; r/E_r = 15.8.$$

This last result is of particular interest, since Batchelor and Reed say of this plantation—

This grove presents the most uniform appearance of any under consideration. The land is practically level, and the soil is apparently uniform in texture. The records show a grouping of several low-yielding trees; yet a field observation gives one the impression that the grove as a whole is remarkably uniform.

Notwithstanding this apparent homogeneity there is a heterogeneity coefficient of over 0.4.

Taking the yields of seedling walnuts in pounds per tree as given in figure 5 of Batchelor and Reed and grouping in a 2 by 2 fold manner, we find

$$r = 0.232 \pm 0.038, \; r/E_r = 6.09.$$

Finally, if the yields in pounds per tree of the Jonathan apple trees mapped by Batchelor and Reed in their figure 6 be treated in a 2 by 2 fold grouping, the coefficient is

$$r = 0.214 \pm 0.043, \; r/E_r = 4.97.$$

Without exception these groves show material values of the heterogeneity coefficients which are statistically significant in comparison with their probable errors throughout.

PHYSICAL AND CHEMICAL BASIS OF THE HETEROGENEITY OF EXPERIMENTAL FIELDS

In foregoing sections it has been shown that when tracts of land are judged by their capacity for crop production the yields are such as to indicate that heterogeneity is a practically universal characteristic of the

fields which may be used for fertilizer tests, variety trials, or any other experimental purpose involving plot yields. In the vast majority of cases the heterogeneity is so great as to leave open to question conclusions drawn from experiments not carried out with all biological precautions and interpreted with due regard to probable errors.

While the actual demonstration of differences in crop yields from one portion of the field to another is the result of final importance from the agronomic standpoint, and while it furnishes all but conclusive evidence that this heterogeneity in yield is due to irregularities in the soil itself, it seems desirable to show that such heterogeneity does actually obtain in the physical and chemical properties of the soil which are determining factors in plant growth.

The desirability of determining the extent to which heterogeneity, in the sense to which the term is used here, obtains in the physical and chemical properties of the soil of experimental fields is emphasized by the following sentences from one of the pioneer papers (*21*) on the variability of soil samples.

> A number of papers have appeared dealing with the variation in the weight of the crop produced over different parts of an apparently uniform field. Such variations reflect the variability of the soil, serving simply as a substratum for the growth of plants, but it is evident that the variations between such measurements as those given do not depend upon the soil as the only variable factor.

At the outset we must recognize that many factors may determine differences in yield. Even if one could secure a tract initially uniform in soil and exposure it is not always possible to be sure that it has all been in the same crop in preceding years. Previous cultures may influence tilth and soil composition by organic remains, by infection with disease-producing organisms, or by differences in the demand of various crops for certain of the plant foods.[1] Such sources of heterogeneity are not readily detected by the eye or by physical or chemical analysis. Even if the experimenter secures a field of sensibly uniform texture, chemical composition, and previous cultural treatment, the uniformity may be readily destroyed in planting or tillage. Rain may interrupt the ploughing, thus exposing the soil of the different portions of the field to air and light for different lengths of time and affecting the physical condition very profoundly. Such sources of error are particularly great in the planting of large experiments. Thus the sources of field heterogeneity can never be fully determined in any case, although individual factors may be demonstrated.

To determine whether an experimental field is heterogeneous with respect to physical or chemical factors, actual measurements of these factors should be made over the field and the heterogeneity coefficient applied. As a first illustration we take a series of soil-moisture

---

[1] These are factors of particular importance in rotation experiments.

determinations uniformly distributed over a plot on a field at the San Antonio Experimental Farm of the Office of Western Irrigation Agriculture.

Hastings (6) has given a condensed account of the soil conditions of the San Antonio region. A map of the experimental farm by Hastings (7, *p. 2*) shows the location of field C3 in which this plot of borings was located [1] and gives meteorological conditions prevailing in 1915, the year in which the borings were made.

Mr. C. S. Scofield kindly informs me that field C3 had been uniformly treated for some time previously and was in apparently uniform condition. It is nearly level but with a gradual slope to the south and east.

The soil has the superficial appearance of uniformity, but we know from experience that the subsoil, which is usually characterized by a high lime content, is in some

1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	21	22	23	24	25	26
27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52
53	54	55	56	57	58	59	60	61	62	63	64	65
66	67	68	69	70	71	72	73	74	75	76	77	78
79	80	81	82	83	84	85	86	87	88	89	90	91
92	93	94	95	96	97	98	99	100				

FIG. 10.—Diagram showing location of sample areas examined for soil moisture in a field at the San Antonio Experimental Farm.

places much closer to the surface than in others. However, from a general agronomic standpoint, this field would be regarded as extremely uniform, and observation of it during the growing season would tend to confirm this view.

Borings were made 6 feet in depth and were sampled at every foot.[2] Figure 10 shows the form of this field.

In order to reduce the 100 sample areas to 2 by 2 fold combinations we have discarded the right file and a portion of one rank, retaining only those which can be grouped into fours as indicated by the cross lines. The percentages of moisture content of these 100 sample areas appear in Table III.[3]

---

[1] The northern border of the sampled area is a line 60 feet south of the north line of the field and parallel to it.

[2] The samples were all taken between March 31 and April 9. During this period there was no rain. Between March 15 and April 10 there were only two rains, one on March 17 of 0.2 inch, the other on March 29 of 0.01 inch. Neither of these was sufficient to affect the soil moisture conditions, since in this region a precipitation of less than 0.25 inch scarcely penetrates the surface-soil mulch. Thus moisture changes during the course of the work can hardly influence the results.

[3] The 12 sample areas which were omitted because of impossibility of combining by fours are starred (*).

TABLE III.—*Moisture content of 100 sample areas of a field at the San Antonio Experimental Farm*

[Expressed in percentages]

Sample area No.	First foot.	Second foot.	Third foot.	Fourth foot.	Fifth foot.	Sixth foot.
I	20. 2	19. I	17. 5	13. I	9. 7	8. 9
2	23. 7	21. 6	19. 8	16. 8	15. 0	15. 9
3	20. 9	20. 5	19. 4	16. I	15. 0	15. I
4	21. 5	20. 3	18. 3	16. 0	15. 4	14. 2
5	23. 3	22. 4	20. 6	19. 3	16. I	15. 8
6	25. 0	24. I	19. 7	17. 6	16. 5	15. 3
7	22. 8	23. 0	20. 8	17. 0	14. 8	14. 8
8	24. 6	24. 3	20. 7	18. 5	15. 8	14. 8
9	25. 6	25. 3	25. 3	25. 5	23. 7	18. 7
10	22. 9	25. 8	26. 0	26. 2	23. 5	18. 6
11	28. 0	30. 4	30. 6	29. 8	26. 8	21. 5
12	25. 2	25. 7	24. 5	26. 8	24. 0	21. 2
13[*]	22. I	22. 0	20. I	20. I	16. 0	14. 9
14	20. 2	19. 7	17. 0	14. 5	II. 4	9. I
15	22. I	21. 3	18. I	14. 6	13. 8	12. 7
16	25. I	21. 2	20. 0	16. 3	15. 5	14. I
17	21. 8	21. 0	19. 2	16. 6	15. I	14. 9
18	23. 4	22. 4	20. 0	16. I	15. 7	15. 4
19	20. 5	20. 8	19. 6	15. 6	13. 5	12. 5
20	24. 0	22. 0	19. 5	15. I	II. 5	9. 4
21	20. 4	20. 7	18. 7	13. 0	8. I	8. 2
22	24. 3	24. 0	21. 0	23. 7	21. 3	15. 2
23	21. 3	22. 2	21. 8	21. 7	20. 5	16. 7
24	24. 3	25. 7	24. 0	22. 4	18. 3	14. 9
25	23. 6	23. 2	24. 4	24. 5	22. 2	18. 4

TABLE III.—*Moisture content of 100 sample areas of a field at the San Antonio Experimental Farm*—Continued

[Expressed in percentages]

Sample area No.	First foot.	Second foot.	Third foot.	Fourth foot.	Fifth foot.	Sixth foot.
26*	24. 2	23. 7	23. 0	20. 7	19. 1	17. 8
27	21. 1	19. 7	18. 7	14. 7	14. 5	17. 7
28	21. 2	19. 6	18. 4	17. 6	15. 2	15. 0
29	21. 2	20. 5	19. 6	18. 9	17. 5	17. 1
30	22. 9	22. 0	19. 9	17. 5	15. 0	14. 8
31	21. 0	20. 7	19. 6	16. 2	14. 0	16. 4
.32	23. 4	21. 6	19. 3	18. 6	16. 8	15. 9
33	22. 2	21. 8	20. 1	16. 6	14. 0	14. 1
34	23. 9	22. 7	20. 4	17. 0	14. 6	14. 2
35	21. 6	20. 9	19. 2	16. 8	15. 3	16. 3
36	21. 4	21. 6	20. 6	20. 0	18. 4	16. 8
37	25. 3	25. 6	25. 6	24. 9	22. 2	17. 9
38	26. 7	29. 2	27. 0	25. 9	23. 0	19. 1
39*	26. 2	29. 8	30. 4	28. 6	26. 1	21. 5
40	21. 8	20. 0	19. 3	15. 7	15. 7	16. 3
41	19. 9	19. 4	19. 0	15. 3	14. 8	14. 9
42	21. 6	20. 0	18. 2	14. 1	15. 5	15. 0
43	19. 6	21. 7	19. 0	14. 7	14. 2	13. 9
44	21. 6	21. 7	19. 4	15. 6	15. 3	15. 4
45	21. 6	20. 5	19. 3	16. 3	7. 5	14. 2
46	22. 6	21. 3	12. 2	16. 2	14. 4	14. 3
47	21. 0	22. 0	19. 3	15. 9	14. 7	15. 0
48	22. 0	21. 5	19. 8	19. 8	14. 7	15. 2
49	22. 7	22. 1	20. 0	19. 5	16. 2	16. 2
50	21. 9	23. 3	21. 0	19. 1	16. 2	16. 5

TABLE III.—*Moisture content of 100 sample areas of a field at the San Antonio Experimental Farm*—Continued

[Expressed in percentages]

Sample area No.	First foot.	Second foot.	Third foot.	Fourth foot.	Fifth foot.	Sixth foot.
51	20. 0	20. 3	19. 0	17. 6	15. 7	17. 2
52*	29. 6	28. 4	27. 3	22. 0	13. 2	16. 2
53	20. 6	19. 8	18. 5	15. 7	15. 9	15. 6
54	21. 2	20. 7	18. 8	15. 1	14. 3	14. 5
55	19. 3	20. 0	18. 9	16. 3	14. 1	14. 9
56	21. 2	20. 8	18. 9	16. 3	14. 1	14. 9
57	22. 1	21. 0	19. 5	15. 7	15. 1	15. 7
58	22. 7	21. 6	19. 7	18. 3	14. 7	15. 8
59	21. 2	21. 0	19. 7	17. 4	15. 2	16. 4
60	23. 2	22. 5	20. 7	19. 1	16. 5	16. 5
61	19. 4	21. 3	19. 7	17. 8	16. 9	17. 2
62	22. 6	21. 3	18. 6	18. 6	15. 0	17. 1
63	21. 3	20. 6	19. 3	17. 5	17. 3	17. 9
64	21. 7	20. 1	18. 9	15. 8	15. 5	16. 6
65*	22. 5	21. 0	20. 2	16. 7	17. 0	20. 8
66	20. 6	21. 2	17. 9	17. 1	15. 8	15. 0
67	18. 9	19. 2	18. 2	15. 0	14. 4	15. 0
68	23. 4	14. 5	19. 0	17. 7	15. 5	15. 5
69	21. 2	20. 4	18. 8	17. 0	14. 0	13. 9
70	21. 4	20. 1	18. 4	17. 0	15. 8	16. 0
71	21. 0	21. 1	18. 9	15. 6	14. 5	15. 3
72	22. 8	21. 4	20. 0	16. 5	15. 5	14. 3
73	21. 9	21. 6	20. 2	16. 2	14. 4	17. 9
74	22. 8	21. 8	20. 3	18. 0	15. 5	17. 9
75	21. 3	22. 8	22. 1	21. 6	17. 7	15. 1

TABLE III.—*Moisture content of 100 sample areas of a field at the San Antonio Experimental Farm*—Continued

[Expressed in percentages]

Sample area No.	First foot.	Second foot.	Third foot.	Fourth foot.	Fifth foot.	Sixth foot.
76	21. 5	22. 0	19. 7	17. 3	17. 2	16. 9
77	21. 4	21. 0	19. 7	15. 8	15. 6	18. 0
78*	22. 4	21. 0	19. 2	17. 0	16. 2	15. 6
79	18. 5	18. 8	18. 4	17. 0	15. 2	15. 1
80	20. 3	19. 6	18. 5	15. 2	15. 0	15. 8
81	20. 3	20. 2	18. 8	16. 5	14. 6	15. 5
82	21. 5	21. 7	18. 6	15. 8	14. 8	14. 1
83	20. 0	20. 4	18. 7	15. 7	16. 3	15. 4
84	20. 3	20. 0	18. 9	17. 5	14. 7	14. 7
85	22. 4	21. 8	21. 4	17. 1	15. 9	14. 8
86	23. 2	22. 0	19. 6	16. 0	15. 7	15. 0
87*	21. 8	21. 6	20. 8	19. 1	17. 2	16. 5
88*	23. 7	21. 8	20. 2	16. 4	16. 9	16. 4
89*	28. 0	21. 6	20. 2	18. 3	17. 0	18. 5
90*	23. 2	21. 7	19. 1	16. 3	16. 3	16. 6
91*	22. 3	22. 9	21. 7	19. 3	18. 5	18. 6
92	20. 2	19. 7	18. 2	17. 4	14. 7	15. 0
93	19. 0	19. 3	18. 5	16. 1	15. 4	15. 9
94	22. 0	20. 4	18. 3	16. 0	14. 9	14. 0
95	21. 5	19. 7	18. 8	14. 9	14. 9	14. 5
96	20. 8	20. 3	18. 7	16. 3	14. 4	15. 4
97	20. 1	19. 5	19. 1	17. 9	15. 0	16. 3
98	22. 6	20. 3	19. 4	15. 3	15. 0	15. 3
99	20. 4	20. 3	18. 6	16. 4	14. 6	14. 5
100*	22. 6	21. 6	19. 4	17. 5	16. 3	15. 0

To determine whether the distribution of soil moisture in these plots is such that it might bring about a correlation between the yields of adjacent plots due to heterogeneity in regard to this physical factor in

the field we have merely to determine the correlations between the percentages of water content of associated plots. These are

Depth.	Correlation.	r/E_r.
First foot.............	0.317±0.065	4.9
Second foot..........	.529± .052	10.2
Third foot...........	.542± .051	10.7
Fourth foot..........	.704± .036	19.4
Fifth foot............	.607± .045	13.4
Sixth foot	.484± .055	8.8

The correlations are of a very substantial order, ranging from 0.317 to 0.704. Notwithstanding the fact that there are only 88 stations upon which the probable errors are based, the constants may in every case be considered significant in comparison with their probable errors.

Thus, notwithstanding the fact that we are dealing with a field only 150 by less than 264 feet,[1] there is a marked and statistically significant heterogeneity in respect to so important a factor in plant growth as soil moisture at each level in the upper 6 feet of soil.

This result seems of very real importance in its relation to the practical phases of plot-test work. It shows beyond all dispute that at least under soil conditions such as are found at the San Antonio Experimental Farm, substratum heterogeneity may be very great at levels of the soil which are ordinarily left entirely out of account in the selection of fields which are to be used for plot tests but which are not below the extensions of the roots of the deeper-penetrating crops and not too deep to serve as reserves of soil moisture for the higher layers of the soil in the case of crops which draw their water from more superficial levels.

It is of some interest to determine whether the correlations at one level in the field may be looked upon as sensibly higher than those at other levels. We have, therefore, determined the differences between the correlations at the different depths. These are given with their probable errors, and in relation to their probable errors, in Table IV.

In the table the positive signs indicate higher correlations at lower levels. Of the 10 possible comparisons between the correlations of the first 5 feet, all but one show greater heterogeneity at the lower levels. The sixth foot seems to be somewhat more homogeneous than the second to the fifth foot. A number of the differences are apparently significant in comparison with their probable errors. Thus there is apparently a real difference in the amount of heterogeneity of this field at different levels. Heterogeneity is least at the surface and greatest at a depth of 4 feet.

The significance of this result will perhaps be apparent at once. A field might be reasonably uniform for the surface foot of soil and hence

---

[1] The total length is 264 feet, but this is reduced by discarding the right file.

fairly well suited to the testing of shallow-rooted crops. Below this it might show a higher degree of heterogeneity. Possibly this heterogeneity of lower-lying strata is the explanation of the large correlations obtained for the yields of neighboring trees in groves planted on apparently uniform soil.

TABLE IV.—*Differences and criteria of trustworthiness of differences in the correlation of adjacent plots in soil moisture determinations at various levels*

Depth.	Second foot.		Third foot.		Fourth foot.		Fifth foot.		Sixth foot.	
	$r.$	$r/Er.$	$r.$	$r/Er.$	$r.$	$r/Er.$	$r.$	$r/Er.$	$r.$	$r/Er.$
First foot..............	+0.212 ± .083	2.56	+0.226 ± .082	2.74	+0.387 ± .074	5.22	+0.291 ± .079	3.68	+0.167 ± .085	1.97
Second foot..........:			+ .013 ± .073	.18	+ .175 ± .063	2.76	+ .078 ± .069	1.14	− .045 ± .076	.60
Third foot..............					+ .161 ± .062	2.58	+ .065 ± .068	.96	− .059 ± .074	.79
Fourth foot...........							− .096 ± .058	1.66	− .220 ± .066	3.34
Fifth foot.............									− .124 ± .071	1.74

We can pursue this question of the relationship between the water content of the plots somewhat further. If the factors which determine the similarity in the moisture contents of the combination plots affect more than a single layer, we should expect a correlation between the contents of the first and second foot, and so on, in the same boring. The possible correlations have been worked out for the first foot and the remaining layers and are as follows:

Depth.	Correlation.	$r/Er.$
First and second feet.................	+0.748 ± 0.032	23.59
First and third feet.................	+ .669 ± .040	16.84
First and fourth feet...............	+ .648 ± .042	15.53
First and fifth feet.................	+ .578 ± .048	12.06
First and sixth feet.................	+ .353 ± .063	5.62

There is a statistically significant and even high correlation between the water content of succcessive levels in the same boring.

When we turn to the problem of chemical heterogeneity, we find that while a number of soil chemists have noted the desirability of considering the variability of the soil in taking samples, the available data suitable for testing the degree of heterogeneity of experimental fields are not extensive.

Kaserer's series of determinations (9) is not sufficiently large or properly distributed over the field to make desirable an attempt to measure heterogeneity. Fortunately Waynick and Sharp (22) have given four excellent series, two for nitrogen and two for carbon, derived from two California fields.

Their samples were taken over a total area of a little more than 1.3 acres on two fields of very different character—a silty clay loam at Davis and a blow sand at Oakley.

The fields were both selected for their apparent uniformity, both being nearly level with no change in the soil mass from one part of the field to another great enough to be detected by the usual field methods. Both fields were practically free from vegetation when selected, and before the samplings were made in March, 1918, all extraneous material had been carefully removed.

Altogether they took 80 samples distributed at 30-foot intervals over the entire area. These samples were arranged in an 8 by 10 fold manner. The original data are given in their Tables 3 and 4. Arranging these in the order of the map of the borings given in their figure 1 and combining in a 2 by 2 fold manner, we derive the following heterogeneity coefficients:

For the silty clay loam at Davis—

$$\text{For carbon,} \quad r = 0.417 \pm 0.063, \ r/E_r = 6.67.$$
$$\text{For nitrogen,} \quad r = .498 \pm .057, \ r/E_r = 8.75.$$

For the blow sand at Oakley—

$$\text{For carbon,} \quad r = 0.317 \pm 0.068, \ r/E_r = 4.65.$$
$$\text{For nitrogen,} \quad r = .230 \pm .072, \ r/E_r = 3.20.$$

All these values are statistically significant in comparison with their probable errors. Although the total number of samples is rather small, they indicate in each case a distinct heterogeneity for these important constituents of the soil. Apparently the two fields differ in their heterogeneity, the coefficients for both carbon and nitrogen being distinctly lower on the blow sand at Oakley than on the silty clay loam at Davis. The average carbon content at Oakley is only 0.444 as compared with 1.109 at Davis, while the nitrogen at Oakley is 0.033 as compared with 0.101 at Davis. Probably greater heterogeneity would be expected on general physical considerations on the silt loam than on the blow sand.

The analysis may profitably be carried one step farther. If these fields are heterogeneous in respect to the soil constituents here under consideration, one might anticipate a correlation between the carbon and the nitrogen content of the samples distributed over these fields. The results are

$$\text{For the Davis loam,} \quad r_{nc} = 0.785 \pm 0.029, \ r/E_r = 27.$$
$$\text{For the Oakley blow sand,} \quad r_{nc} = .744 \pm .034, \ r/E_r = 22.$$

Both constants are large. They show that the field is not merely heterogeneous but that portions which are high in nitrogen are high also in carbon and vice versa.

Waynick (*21*) has given a series of 81 determinations of nitrification in samples of soil drawn from a field on the University of California farm at Davis.

The field had been planted to corn in 1914, to Sudan grass in 1915, and to grain sorghum in 1916. In 1917 it had lain fallow and was without vegetation when the samples were taken October 20.

The particular area chosen was apparently as uniform as one could well find, being level, of uniform texture and color, and free from small local depression of any kind.

These samples were taken on eight radii of a circle 100 feet in diameter. The samples were separated by a radial distance of 5 feet. Disregarding the one central sample, we may group the remainder by twos in order to determine whether there is a correlation between adjacent samples. The coefficients thus obtained will, of course, not be comparable with those deduced for cases in which the yields or soil samples were uniformly distributed over the field. They will, however, serve to indicate whether or not this field is heterogeneous in the sense that differences prevailed sufficiently large to influence the properties of adjacent samples in a manner to make them more similar than pairs of samples taken at random over the field. His samples were drawn in two series—the first from the superficial 6 inches, the second from the deeper-lying level, 6 to 24 inches.

Waynick's Table 1 gives the residual nitrate in soil as sampled. From it we deduce

For the upper 6 inches, $r = 0.404 \pm 0.063$, $r/E_r = 6.4$.
For the subsoil, $r = .596 \pm .049$, $r/E_r = 12.2$.

Table 2 gives the nitrate produced from the soil's own nitrogen after 28 days' incubation. We deduce

For the upper 6 inches, $r = 0.065 \pm 0.075$, $r/E_r = 0.86$.
For the subsoil, $r = .059 \pm .075$, $r/E_r = .79$.

Table 3 shows the nitrate produced from 0.2 gm. of ammonium sulphate in 100 gm. of soil. The correlation coefficients are

For the upper 6 inches, $r = 0.298 \pm 0.069$, $r/E_r = 4.34$.
For the subsoil, $r = .351 \pm .066$, $r/E_r = 5.31$.

Finally, Table 4 shows the nitrate produced from 0.2 gm. of blood in 100 gm. of soil. The results in this case are

For the upper 6 inches, $r = 0.120 \pm 0.074$, $r/E_r = 1.62$.
For the subsoil, $r = .297 \pm .069$, $r/E_r = 4.32$.

The coefficients show that for both the upper and lower soil layers there is a correlation of about medium value between adjacent samples for the residual nitrate in the soil. These coefficients are unquestionably significant in comparison with their probable errors.

While the coefficients for nitrogen produced from soil nitrogen after incubation are both positive in sign, neither can be considered statistically trustworthy in comparison with its probable error. When nitrogen is added to the soil, in the form of either ammonium sulphate or of blood, the correlations between the nitrogen produced on incubation are larger. All are positive in sign, and three of the four may be reasonably considered statistically significant.

Thus it is clear that this plot, only 100 feet in diameter, shows distinct heterogeneity in residual nitrate and in the amount of nitrification occurring on incubation after the addition of nitrogen.

SUMMARY AND CONCLUSIONS

The purpose of this paper, which is one of a series on the statistical phases of the problem of plot tests, is to show the extent to which the heterogeneity of experimental fields may influence plot yields.

By heterogeneity we understand differences in capacity for crop production throughout the field of such a magnitude as to influence in like manner, but not necessarily to like degree, the yield of adjacent small plots. Thus, variability of plot yields does not necessarily indicate the heterogeneity of the fields upon which tests are made but may be due to other factors.

Heterogeneity is measured by a coefficient which shows the degree of correlations between the yields of associated ultimate plots, grouped in combination plots.

This coefficient has been determined for a relatively large series of experimental fields widely distributed throughout the world and planted to a considerable variety of crops, for which a number of different kinds of yields have been measured. The results show that in every field the irregularities of the substratum have been sufficient to influence, and often profoundly, the experimental results.

It might be objected that by chance, or otherwise, the illustrations are not typical of what ordinarily occurs in plot cultures. But the series considered practically exhaust the available data for such purposes. Furthermore the records are in large part drawn from the writings of those who are recognized authorities in agricultural experimentation and who have given their assurance of the suitability of the fields upon which the tests were made.

For example, Mercer and Hall (*15*) state the purpose of their research to be—

to estimate the variations in the yield of various sized plots of ordinary field crops which had been subjected to no special treatment and appealed to the eye sensibly uniform.

Their mangolds—

looked a uniform and fairly heavy crop for the season and soil,

while in their wheat field—

a very uniform area was selected.

The data of Larsen were drawn from an experiment—

auf einer dem Auge sehr gleichmassig erscheinenden, 3 Jahre alten Timotheegraswiese.

Montgomery's data were secured from a plot of land only 77 by 88 feet in size, which had been sown continuously to Turkey wheat for three years—

and was of about average uniformity and fertility.

Coombs and Grantham selected a field on which—

the crop was extremely regular as judged before the cutting and it had not been subjected to any attack of borer or any devastation of rats or birds.

Lyon's potato field was selected from—

a piece of apparently uniform land.

Mr. C. S. Scofield kindly informs us that the Huntley tract was selected for apparent uniformity and that prior to the calculation of the constants presented in this paper there was no reason, from general observation, to suspect irregularities in the field. Batchelor and Reed have assured me that their orchards are to all appearances uncommonly uniform. Kiesselbach emphasizes the apparent uniformity of his oat field.

Nothing could more emphasize the need of a scientific criterion for substratum homogeneity than the fact that correlations between the yields of adjacent plots ranging from $r = +0.020$ to $r = +0.830$ can be deduced from the data of fields which have passed the trained eyes of agricultural experimenters as satisfactorily uniform.

A second phase of this investigation has been to ascertain whether the physical or chemical requisites for plant growth are so distributed over experimental fields that they may be reasonably looked upon as the source of the demonstrated heterogeneity in yield.

The heterogeneity coefficients for percentage of water content for the upper 6 feet on the Experimental Farm of the Office of Western Irrigation Agriculture at San Antonio, Tex., range from $+0.32$ to $+0.70$ and are statistically significant for each of the 6 upper feet of soil. Heterogeneity is least at the surface and greatest at a depth of 4 feet. The surface layer of soil might, therefore, be apparently uniform in water content while underlying layers might differ greatly from one part of the field to another. This may be the explanation of the correlation between the yields of adjacent trees in groves planted in an apparently uniform locality.

Analysis of the data of Waynick and Sharp shows that there is a correlation of from $+0.23$ to $+0.50$ between adjacent borings for so important soil constituents as nitrogen and carbon. The correlation

between nitrogen content and carbon content of samples from two different soils is of the order + 0.75.

It is interesting to note that these coefficients for water content and for chemical composition of the soil are of about the same order as those found for crop yields. While these results do not prove that the heterogeneity of experimental fields in their capacity for crop production is directly due to these and other physical and chemical factors, there can be little doubt that this is actually the case.

The references here made to the existence of significant heterogeneity in fields passed by agricultural experts as satisfactorily uniform must not be interpreted as a criticism of the work of these investigators. There is, indeed, every evidence of care and thoroughness. The result merely illustrates the inadequacy of personal judgment concerning the uniformity in physical characters or in crop-producing capacity of fields under consideration for experimental work.

The demonstration that the fields upon which plot tests have been carried out in the past are practically without exception so heteregeneous as to influence profoundly the yields of the plots emphasizes the necessity for greater care in agronomic technic and more extensive use of the statistical method in the analysis of the data of plot trials if they are to be of value in the solution of agricultural problems.

To other phases of the problem we shall return in subsequent papers.

LITERATURE CITED

(1) BATCHELOR, L. D., and REED, H. S.
    1918. RELATIONSHIP OF THE VARIABILITY OF YIELDS OF FRUIT TREES TO THE ACCURACY OF FIELD TRIALS. *In* Jour. Agr. Research, v. 12, no. 5, p. 245–283, 11 fig. Literature cited, p. 282–283.

(2) COOMBS, G. E., and GRANTHAM, J.
    1916. FIELD EXPERIMENTS AND THE INTERPRETATION OF THEIR RESULTS. *In* Agr. Bul. Fed. Malay States, v. 4, no. 7, p. 206–216, 1 fig.

(3) HARRIS, J. Arthur.
    1913. ON THE CALCULATION OF INTRA-CLASS AND INTER-CLASS COEFFICIENTS OF CORRELATION FROM CLASS MOMENTS WHEN THE NUMBER OF POSSIBLE COMBINATIONS IS LARGE. *In* Biometrika, v. 9, pt. 3/4, p. 446–472.

(4) ——
    1914. ON SPURIOUS VALUES OF INTRA-CLASS CORRELATION COEFFICIENTS ARISING FROM DISORDERLY DIFFERENTIATION WITHIN THE CLASSES. *In* Biometrika, v. 10, pt. 2/3, p. 412–416.

(5) ——
    1915. ON A CRITERION OF SUBSTRATUM HOMOGENEITY (OR HETEROGENEITY) IN FIELD EXPERIMENTS. *In* Amer. Nat., v. 49, no. 583, p. 430–454.

(6) HASTINGS, S. H.
    1916. THE WORK OF THE SAN ANTONIO EXPERIMENT FARM IN 1915. U. S. Dept. Agr. Bur. Plant Indus. West. Irrig. Agr. 10 (Misc. Pub.), 17 p., 1 fig.

(7) —— and BLAIR, R. E.
    1915. HORTICULTURAL EXPERIMENTS AT THE SAN ANTONIO FIELD STATION, SOUTHERN TEXAS. U. S. Dept. Agr. Bul. 162, 26 p., 8 fig.

(8) HOLTSMARK, G., and LARSEN, B. R.
    1906. ÜBER DIE FEHLER, WELCHE BEI FELDVERSUCHEN DURCH DIE UNGLEI-
        CHARTIGKEIT DES BODENS BEDINGT WERDEN. *In* Landw. Vers. Stat.,
        Bd. 65, Heft 1/2, p. 1–22.

(9) KASERER, Hermann.
    1910. BEITRAG ZUR FRAGE DER FESTELLUNG DES NÄHRSTOFFGEHALTES EINER
        ACKERPARZELLE. *In* Ztschr. Landw. Versuchsw. Oesterr., Jahrg. 13,
        Heft 8, p. 742–747, 1 fig.

(10) KIESSELBACH, T. A.
    1918. STUDIES CONCERNING THE ELIMINATION OF EXPERIMENTAL ERROR IN
        COMPARATIVE CROP TESTS. Nebr. Agr. Exp. Sta. Research Bul. 13,
        95 p., 20 fig.

(11) ————
    1919. EXPERIMENTAL ERROR IN FIELD TRIALS. *In* Jour. Amer. Soc. Agron.,
        v. 11, no. 6, p. 235–241.

(12) LEHMANN, A.
    1901–1909. SECOND TO NINTH ANNUAL REPORTS OF THE AGRICULTURAL CHEMIST.
        Department of agriculture, Mysore State [India]. 1900/01–1907/08.

(13) LOVE, H. H.
    1919. THE EXPERIMENTAL ERROR IN FIELD TRIALS. *In* Jour. Amer. Soc.
        Agron., v. 11, no. 5, p. 212–216.

(14) LYON, T. L.
    1912. SOME EXPERIMENTS TO ESTIMATE ERROR IN FIELD PLAT TESTS. *In* Proc.
        Amer. Soc. Agron., v. 3, p. 89–114, 5 fig.

(15) MERCER, W. B., and HALL, A. D.
    1911. THE EXPERIMENTAL ERROR OF FIELD TRIALS. *In* Jour. Agr. Sci., v. 4,
        pt. 2, p. 107–132, 10 fig.

(16) MONTGOMERY, E. G.
    1912. VARIATION IN YIELD AND METHOD OF ARRANGING PLOTS TO SECURE COM-
        PARATIVE RESULTS. *In* Nebr. Agr. Sta. 25th Ann. Rpt., [1911]/12,
        p. 164–180, 4 fig.

(17) ————
    1913. EXPERIMENTS IN WHEAT BREEDING: EXPERIMENTAL ERROR IN THE
        NURSERY AND VARIATION IN NITROGEN AND YIELD. U. S. Dept. Agr.
        Bur. Plant Indus. Bul. 269, 61 p., 22 fig., 4 pl.

(18) SMITH, Louie H.
    1910. PLOT ARRANGEMENT FOR VARIETY EXPERIMENTS WITH CORN. *In* Proc.
        Amer. Soc. Agron., v. 1, 1907/09, p. 84–89.

(19) STEWART, F. C.
    1919. MISSING HILLS IN POTATO FIELDS: THEIR EFFECT UPON THE YIELD. N. Y.
        Agr. Exp. Sta. Bul. 459, p. 45–69, 3 fig.

(20) STOCKBERGER, W. W.
    1916. RELATIVE PRECISION OF FORMULAE FOR CALCULATING NORMAL PLOT
        YIELDS. *In* Jour. Amer. Soc. Agron., v. 8, no. 3, p. 167–175.

(21) WAYNICK, D. D.
    1918. VARIABILITY IN SOILS AND ITS SIGNIFICANCE TO PAST AND FUTURE SOIL
        INVESTIGATIONS. I. A STATISTICAL STUDY OF NITRIFICATION IN SOILS.
        Univ. Cal. Pub. Agr. Sci., v. 3, no. 9, p. 243–270, 2 fig.

(22) ———— and SHARP, L. T.
    1919. VARIABILITY IN SOILS AND ITS SIGNIFICANCE TO PAST AND FUTURE SOIL
        INVESTIGATIONS. II. VARIATIONS IN NITROGEN AND CARBON IN FIELD
        SOILS AND THEIR RELATION TO THE ACCURACY OF FIELD TRIALS. Univ.
        Cal. Pub. Agr. Sci., v. 4, no. 5, p. 121–139, 1 fig.

PERMANENCE OF DIFFERENCES IN THE PLOTS OF AN EXPERIMENTAL FIELD

BY

J. ARTHUR HARRIS AND C. S. SCOFIELD

PERMANENCE OF DIFFERENCES IN THE PLOTS OF AN EXPERIMENTAL FIELD

By J. Arthur Harris, *Investigator, Station for Experimental Evolution, Cold Spring Harbor, N. Y., and Collaborator, Office of Western Irrigation Agriculture,* and C. S. Scofield, *Agriculturist in Charge, Office of Western Irrigation Agriculture, Bureau of Plant Industry, United States Department of Agriculture*

I.—INTRODUCTION

Agronomists have long recognized the fact that the plots of an experimental field may differ considerably among themselves. This variability is the source of the greatest difficulty in the interpretation of comparative cultures. A recent analysis (*3*)[1] of the available data by adequate biometric formulae (*1*) has shown that heterogeneity is a practically universal characteristic of experimental fields and that it must be considered in the interpretation of the results of all plot tests.

With the demonstration of this characteristic of experimental areas the questions naturally arise: Are the differences between plots transient or are they relatively permanent from year to year? Do these differences tend to increase or to decrease with cultivation?

Presumably the differences which obtain in the soil of an experimental field are in part permanent and in part transient. Lyon (*5*) suggested that tillage and other factors will change the plots so that the results will not be comparable from year to year. Unfortunately he does not present data to show to what extent this may be true. He gives a series of yields for successive years on the same plots, which measured 33 by 66 feet or $1/20$ of an acre in area, at the Nebraska Agricultural Experiment Station and shows that the rank of the yield of these plots differs greatly from year to year. Thus he concludes that if they differ among themselves in their capacity for crop production this difference is very little constant from year to year.

Smith (*6*) took advantage of the breaking up of a piece of land which had lain 16 years in pasture to investigate the effect of cultivation on the uniformity of a series of plots. Any influence of 1 or 2 years preceding cultures on the variation or correlation of yields should, he assumed, be apparent in the statistical constants deduced from these

---

[1] Reference is made by number (italic) to "Literature cited," p. 356.

data. He gives a table which indicates that there is such a change. He says:

> It is noticeable that the variability as measured by the standard deviation becomes less in each succeeding year. This suggests the question as to whether continued cropping might not tend to induce uniformity. The records of a few of these plots which were continued in corn for three years longer do not support such a conclusion.

It must be noted that in Smith's experiments seasonal conditions varied greatly from year to year. Thus 1895, which was exceedingly dry and also cool in the early part of the season, was highly unfavorable. The two following years were unusually favorable for corn. As a result the yields were, respectively, 31.6, 91.6, and 71.4 bushels per acre in the three years.

Lehmann in his work at the experimental farm near Bangalore attempted to use the experience of previous years in the standardization of experimental plots. His data will be considered in some detail below.

II.—METHODS AND RESULTS

The permanency of the differentiation of plots in their capacity for crop production may be measured in terms of correlation. If the plots of a field differ among themselves in a more or less permanent way there will, with reasonably uniform climatic conditions, be a correlation between the yields of the plots of a series in two or more successive years—in short, an interannual correlation (2).

The problem of the correlations between the yields of identical plots in different years is one of very great interest. If this correlation be high it should be possible to standardize a field of plots by one or more sowings to the same variety. A chief difficulty in the standardization of the field by the carrying out of a preliminary test in which the productive capacities of the plots are determined once and for all lies in the fact that the factors which determine yield are in part edaphic— that is, pertaining to soil conditions—and in part meteorological. For example, in a very dry year sections of a field which are lower may produce the heaviest crops because adequate moisture is longer retained in these places. In a wet year the case may be just the reverse, for the crops in the lower-lying portions may be too wet for the best plant growth. Thus, it is quite possible that in cases in which there is a profound influence of environmental factors there may be a negative correlation between the yield of the same plots in different years.

It is conceivable, therefore, that the interannual correlation for yield per plot may range from negative to positive values, zero correlation being found in cases in which edaphic and meteorological factors exactly counterbalance each other in their influence upon the yield of the plots of a heterogeneous field.

A.—PUBLISHED DATA

Unfortunately few data are available for analysis from the literature.

Lehmann has given (*4, p. 6*) yields of paddy on the 17 plots of ranges B and C, respectively, of the wet tract of the Experimental Farm at Hebbel. Grouping the yields for the two ranges, we find for the correlations between the yields of the same plots in the two years 1905 and 1906:

$$\text{Range B, } r = 0.834 \pm 0.050, \; r/Er = 16.7.$$
$$\text{Range C, } r = 0.799 \pm 0.059, \; r/Er = 13.5.$$

Stockberger (*7*) gives data for the extremes of a series of hill yields for hops. The interannual correlations deduced from these data have been shown (*2*) to be as follows:

Years.	Lowest hills.	Highest hills.
1909 and 1910	0. 29±0. 17	0. 59±0. 13
1910 and 1911	.55± .13	.52± .14
1909 and 1911	.43± .15	.30± .18

Stockberger has also given (*8*) the yields for 30 rows, each 210 feet in length, from hop fields of several hundreds of acres in the Sacramento Valley of California:

The plants in these rows averaged well in number and uniformity of growth with the plants on several hundreds of acres of hops in the midst of which the experimental area was located.

Data are available for the years 1909 to 1914. Calculating the correlation between the yields in the different years, we have the results set forth in Table I. It appears that with one single exception the constants are positive throughout. In general they are significant in comparison with their probable errors, indicating a superiority in a subsequent year if a superiority is shown in a given year.

The constants in the table are arranged in a way to show the change in the coefficient of correlation as the years become more widely separated in time. Thus, in the case of the correlation for the 1909 yields, the constant for "first and second" is that showing the relationship between the 1909 and 1910 yields, while "first and third" indicates the constant measuring the relationship between the yields of 1909 and 1911. Similarly, in the series of coefficients for 1910 "first and second" designates the correlation between 1910 and 1911, etc.

For the series beginning with 1909 we note a marked decrease in the magnitude of the constants as the yields correlated become more widely separated in time. The same is true for the series beginning with 1910. The other series are more irregular.

Table I.—*Interannual correlations for yield of hops*

Beginning of series.	First and second years.	First and third years.	First and fourth years.	First and fifth years.	First and sixth years.
1909	+0.768±0.051	+0.622±0.075	+0.380±0.105	+0.259±0.115	+0.061±0.123
1910	+ .577± .082	+ .447± .099	+ .451± .098	+ .274± .114	
1911	+ .062± .123	+ .313± .111	− .126± .121		
1912	+ .311± .111	+ .705± .062			
1913	+ .597± .079				

The most reasonable explanation of the higher correlation of more closely associated years is that both field conditions and the productiveness of the individual vines change more or less as time goes on. The result of such changes would be a lower correlation between the yields of periods more widely separated in time.

The data for the dry-land experiments in Mysore State have been discussed elsewhere (*3*) in relation to the problem of field heterogeneity. It was shown there that in two dry years the field showed marked hetereogeneity, but that in one unusually wet season there was marked abnormality of yield with little correlation between the yields of adjacent plots.

It seems of unusual interest, therefore, to determine to what extent the differences between these plots are permanent from year to year. Correlating between the yields of ragi, we find the following correlation coefficients for the whole series of 105 plots for which data are available.

Years.	Grain.	Straw.	Total.
1905 and 1906	0. 591±0. 043	0. 777±0. 026	0. 757±0. 028
1905 and 1907	. 693± .034	.855± .018	.852± .018
1906 and 1907	.450± .052	.678± .036	.610± .041

The correlations are of very substantial order, and without exception they are clearly significant in comparison with their probable errors. They show that the differences in the plots are to a high degree persistent during the three years of this experiment.

For grain, straw, and total yield the correlations between the yield for 1905 and 1907 are higher than those for 1905 and 1906 or for 1906 and 1907. If there were a progressive change in the field one might have expected that the correlations would be higher between consecutive years. Apparently the influence of the abnormal conditions of 1906 has been to lower the correlations for this year.

The results show that the capacity for production is to a high degree persistent from year to year, notwithstanding great diversity in meteorological conditions.

A series of records of unusual interest is provided by Smith (*6*) for yields of corn in three successive years, 1895, 1896, 1897. It has been

shown elsewhere (*3*) that this field, which had lain in grass for many years, is highly heterogeneous, showing correlations between adjacent plots of $r = 0.61$ to $r = 0.83$.

The conditions for corn production differed very greatly in the three years. Thus the constants for yield were:

Year.	Mean.	Standard deviation.	Coefficient of variation.
1895...	31.7	7.91	25.0
1896...	91.6	10.64	·11.6
1897...	71.4	6.27	8.8

Yield is over twice as heavy in the second and third years as in the first. The variability in yield as measured by the coefficient of variation is far lower in the second and third years than in the first.

Computing the correlations between the yields for the three years, we have the following results:

For 1895 and 1896, $r = -0.354 \pm 0.054$, $r/E_r = 6.6$.

For 1895 and 1897, $r = -0.221 \pm 0.059$, $r/E_r = 3.8$.

For 1896 and 1897, $r = +0.818 \pm 0.020$, $r/E_r = 40.1$.

There is a negative correlation for 1895 and 1896 and for 1895 and 1897 but a high positive correlation for 1896 and 1897. Thus the plots which were better in the highly unfavorable year 1895 were poorer in the two favorable years 1896 and 1897. Plots which were better in the favorable year 1896 were also better in the favorable year 1897.

B.—THE HUNTLEY UNIFORM CROPPING EXPERIMENT

The most extensive series of records available is that for a uniform cropping experiment conducted for the past several years at the Field Station of the Office of Western Irrigation Agriculture, at Huntley, Mont.

The Huntley field lies in the Yellowstone Valley on land having a very slight and uniform slope to the north. The detailed history of the field prior to 1910 is not known definitely. It was probably first broken from the native prairie sod in the spring of 1908. In 1909 it was planted to sugar beets, but the crop was destroyed in the late summer. It came under experimental control in 1910, when the major portion of it was sown to oats and yielded a crop of 66 bushels per acre. In that season a small tract in the northeast corner of the field was used as a machinery park or stack yard and was not put into crop. This tract occupied about two-thirds of the length of the first five plots in series II. It is possible that this difference of treatment in 1910 may have been reflected in the crop yields of 1911, but it seems hardly probable that any material effects could have persisted longer.

In the spring of 1911 this field was laid out into 46 plots, each measuring 23½ by 317 feet and containing 0.17 acre, arranged in two parallel series of 23 plots each. The two series of plots were separated merely by a temporary irrigation ditch. In 1911 it was planted to sugar beets, and in the spring of 1912 it was seeded to alfalfa, and one cutting was harvested that year. This stand remained on the ground during 1913 and 1914, when the entire field was fall-plowed. In 1913 three cuttings were made, but the third cutting was lost in a heavy wind which scattered and mixed the crop before weighings from the various plots could be made. The first cutting, designated as alfalfa I, was made on plots one-half the original size. The second cutting was harvested from plots one-quarter the original size. The first and second cuttings in 1914 were weighed for plots one-quarter the original size—that is, 0.0425-acre plots—while the third cutting was recorded for plots one-third the original size. These furnish the data for alfalfa I, II, and III for 1914. Total yields for the first and second cuttings in 1913 and 1914 and for the first, second, and third cuttings in 1914 are also considered.

In 1915 and 1916 ear corn was grown. In 1917 [1] the fields were planted to oats, and records were made of grain, straw, and total yield. In 1918 silage corn was grown. In 1919 the land produced a crop of barley.

It has been the practice each season to treat the whole field as a unit until harvest time, when the plot boundaries are established in order to measure the crop yields. In other words, all cultural operations, including irrigation, are carried out on a field scale and uniformly throughout the field. No manuring has so far been attempted. An effort has been made to avoid any artificial causes of heterogeneity.

The crop yields each year have been satisfactory—that is, they have not been abnormal—as is shown in Table II, where the mean yields per plot and per acre are set down. Fortunately, this experiment has also escaped injury from insect pests, plant diseases, and storms, which so often interfere with the success of long-term field experimentation.

[1] Because of other activities the plots could not be harvested in halves and quarters in 1917–1919.

TABLE II.—*Mean yields of the Huntley uniform cropping experiment*

Crop.	Number of pounds per plot.	Number of tons or bushels per acre.
1911, sugar beets....................................	4,179.00	12.29
1912, total alfalfa..................................	356.54	1.04
1913, alfalfa I.....................................	541.41	1.59
1913, alfalfa II....................................	483.26	1.42
1913, alfalfa I and II..............................	1,024.67	3.01
1914, alfalfa I.....................................	489.13	1.44
1914, alfalfa II....................................	499.34	1.47
1914, alfalfa I and II..............................	988.47	2.91
1914, alfalfa III...................................	471.95	1.38
1914, alfalfa I to III...............................	1,460.43	4.29
1915, ear corn.....................................	522.58	52.7
1916, ear corn.....................................	396.15	41.6
1917, oat grain....................................	555.80	102.1
1917, oat straw....................................	521.54	1.53
1917, total yield...................................	1,077.34	3.16
1918, silage corn...................................	3,175.43	9.34
1919, barley grain..................................	358.19	43.8
1919, barley straw..................................	230.50	.67
1919, total yield...................................	588.69	1.73

The data furnished by this series of records are of particular value, since (*a*) they are based on irrigated plots and (*b*) it is possible to compare the correlations between the same crop and different crops in the different years.

The correlations between the yields of the various crops in the different years may be considered in three series.

(1) The first comprises the yields for the whole plots. In this series we determine the correlation between the crop produced on the 46 plots in one year and that produced on the same 46 plots in another year.

(2) In the study of certain crops the plots were divided into two sub-plots, and we may determine the relationship between yield of individual subplots in different years. Then the number of observations is twice what it was in the preceding correlation, that is, $N = 92$ instead of 46.

(3) Finally, in a more limited series of cases the 46 original plots were harvested in 4 subplots each, thus increasing the number of units which may be entered in the correlation tables to 184.

The data for determining the correlations between yields of various crops for the 46 whole plots are given in Table III. The data for half plots and quarter plots may be obtained from the diagrams in an earlier paper by Harris (*3*) on the practical universality of field heterogeneity as a factor affecting plot yields. The correlation coefficients and their probable errors for whole plots are shown in Table IV.

TABLE III.—*Yield of plots of field B at the Huntley (Mont.) Field Station* [a]

Plot No.	1911, sugar beets.	1912, total alfalfa.	1913, alfalfa I.	1913, alfalfa II.	1913, alfalfa I and II.	1914, alfalfa I.	1914, alfalfa II.	1914, alfalfa I and II.	1914, alfalfa III.	1914, alfalfa I to III.
II, 1	12.78	260	595	600	1,195	585	550	1,135	580	1,715
2	12.70	395	530	560	1,090	610	605	1,215	605	1,820
3	10.04	397	640	630	1,270	605	690	1,295	595	1,890
4	10.35	435	640	650	1,290	640	660	1,300	610	1,910
5	9.33	442	625	540	1,165	590	700	1,290	595	1,885
6	9.40	419	625	595	1,220	645	735	1,380	510	1,890
7	11.53	438	640	575	1,215	625	775	1,400	500	1,900
8	12.40	410	555	570	1,125	555	725	1,280	500	1,780
9	10.28	418	570	470	1,040	590	615	1,205	475	1,680
10	11.81	393	540	470	1,010	545	505	1,050	450	1,500
11	13.99	405	585	435	1,020	580	430	1,010	400	1,410
12	12.28	435	530	450	980	555	425	980	445	1,425
13	11.91	385	565	485	1,050	465	445	910	455	1,365
14	11.42	395	555	510	1,065	540	480	1,020	485	1,505
15	12.28	405	655	565	1,220	535	515	1,050	540	1,590
16	13.76	305	650	475	1,125	545	440	985	475	1,460
17	11.73	312	590	435	1,025	540	435	975	460	1,435
18	12.49	290	635	425	1,060	540	525	1,065	455	1,520
19	15.55	315	635	455	1,090	545	490	1,035	465	1,500
20	11.93	310	605	440	1,045	540	505	1,045	475	1,520
21	13.52	330	625	455	1,080	580	535	1,115	510	1,625
22	14.36	325	625	500	1,125	610	525	1,135	520	1,655
23	16.81	310	590	425	1,015	490	445	935	380	1,315
III, 1	13.93	405	535	425	960	420	470	890	645	1,535
2	13.04	350	470	430	900	430	395	825	520	1,345
3	10.55	400	510	405	915	395	435	830	495	1,325
4	11.63	435	475	425	900	440	450	890	440	1,330
5	10.56	350	460	445	905	435	420	855	455	1,310
6	10.00	365	510	510	1,020	430	375	805	415	1,220
7	10.54	390	500	440	940	410	400	810	445	1,255
8	10.00	325	455	425	880	415	380	795	415	1,210
9	8.85	360	490	375	865	425	410	835	385	1,220
10	10.48	360	440	415	855	365	390	755	360	1,115
11	12.61	335	485	390	875	360	420	780	385	1,165
12	11.22	350	470	400	870	360	430	790	370	1,160
13	12.08	370	500	450	950	390	505	895	435	1,330
14	11.91	255	470	485	955	370	495	865	425	1,290
15	12.65	370	485	455	940	380	470	850	455	1,305
16	11.71	325	460	440	900	360	455	815	410	1,225
17	12.19	280	460	445	905	395	425	820	430	1,250
18	12.62	280	430	500	930	395	425	820	385	1,205
19	13.45	320	480	515	995	450	515	965	475	1,440
20	15.60	275	520	565	1,085	435	480	915	445	1,360
21	16.25	290	460	510	970	435	480	915	465	1,380
22	14.70	345	530	535	1,065	475	515	990	495	1,485
23	16.52	337	505	530	1,035	475	475	950	475	1,425

[a] All yields are given in pounds per plot with the exception of that for sugar beets, which is given in tons per acre.

TABLE III.—*Yield of plots of field B at the Huntley (Mont.) Field Station* [a]—Con.

Plot No.	1915, ear corn.	1916, ear corn.	1917, oat grain.	1917, oat straw.	1917, total yield.	1918, silage corn.	1919, barley grain.	1919, barley straw.	1919, total yield.
II, 1	556	513	580	574	1,154	3,655	392	288	680
2	598	514	593	631	1,224	3,285	349	251	600
3	526	481	606	588	1,194	3,290	377	253	630
4	558	495	598	414	1,012	3,390	352	218	570
5	509	487	614	590	1,204	3,570	414	246	660
6	521	450	596	584	1,180	3,240	426	264	690
7	499	489	572	458	1,030	3,005	463	262	725
8	502	441	574	524	1,098	3,010	424	276	700
9	515	434	553	495	1,048	3,060	425	265	690
10	513	415	614	606	1,220	2,885	422	298	720
11	524	399	574	578	1,152	2,955	386	224	610
12	507	379	548	510	1,058	3,055	365	240	605
13	528	376	537	523	1,060	3,125	350	220	570
14	507	372	540	522	1,062	3,210	368	222	590
15	511	398	518	616	1,134	3,155	344	191	535
16	524	409	564	570	1,134	2,870	351	204	555
17	520	389	499	481	980	2,950	333	127	460
18	479	408	538	518	1,056	3,235	309	241	550
19	455	404	637	605	1,242	3,330	313	177	490
20	489	383	579	497	1,076	3,150	304	221	525
21	519	455	567	513	1,080	3,180	316	229	545
22	573	413	553	477	1,030	3,075	306	199	505
23	578	414	509	391	900	3,375	288	257	545
III, 1	545	404	563	547	1,110	3,685	332	238	570
2	552	376	560	522	1,082	3,365	362	218	580
3	504	337	511	511	1,022	3,315	375	260	635
4	547	318	523	497	1,020	3,170	342	183	525
5	544	338	532	516	1,048	3,240	416	284	700
6	533	312	536	552	1,088	3,290	460	250	710
7	505	311	538	544	1,082	2,855	410	330	740
8	519	345	552	556	1,108	2,905	400	260	660
9	513	353	515	535	1,050	2,965	400	260	660
10	509	337	521	545	1,066	2,760	386	274	660
11	493	322	473	479	952	2,640	403	262	665
12	496	357	520	462	982	2,850	305	255	560
13	503	343	645	377	1,022	2,880	296	199	495
14	496	333	525	469	994	3,190	290	130	420
15	518	360	557	485	1,042	3,100	301	174	475
16	499	372	578	504	1,082	2,975	335	185	520
17	483	353	549	515	1,064	2,995	317	188	505
18	469	367	563	517	1,080	3,315	320	190	580
19	477	410	562	512	1,074	3,540	293	187	400
20	490	407	561	481	1,042	3,280	323	177	590
21	551	426	486	456	942	3,370	331	259	580
22	628	423	573	571	1,144	3,625	362	218	510
23	654	401	561	573	1,134	3,705	341	249	590

[a] All yields are given in pounds per plot with the exception of that for sugar beets, which is given in tons per acre.

From the series of correlations as a whole it appears that of the 152 coefficients showing the relationship between crop yields in different years, 133 are positive while only 19 are negative in sign. If the differences in capacity for crop production demonstrated in different years were due to purely transient causes, one would expect to find an approximately equal number of positive and negative correlations with the general average value sensibly zero. Instead we find the proportion of 133 to 19. This is a deviation from the ratio 76 to 76, which one might expect on the assumption that there is no correlation between the yields of plots in a series of years, of

$$57 \pm 0.6745 \sqrt{152 \times 5 \times 0.5} = 57 \pm 4.16.$$

The deviation from equality is 13.7 times as large as its probable error and is unquestionably significant.

If we consider that coefficients which are 2.5 times or more as large as their probable errors represent statistically significant interrelationships, we find that of the 82 relationships which may be regarded as falling in this class 78 are positive whereas only 4 are negative in sign.

Averaging the values of the coefficients considered in Table IV, we note that the average for the 133 positive values is $+0.3346$, whereas that for the 19 negative values is -0.1475. Taking the constants altogether, the average value is $+0.2743$.

There is, therefore, an overwhelming body of evidence to show that plots, even of the small size and the apparent uniformity of those of the Huntley Station, which yield higher in one year will yield higher persistently throughout a series of years.

It is now desirable to determine whether the same relationships hold when these plots are divided into smaller subplots. It is possible to subdivide a number of the plots into 2 subplots, each one-half the original size. Correlations may be determined for the 92 yields of these half plots in the same manner as for the total yields on the 46 original plots. The results appear in Table V.

The constants are positive throughout. In general, they are statistically significant in comparison with their probable errors. As a matter of fact, only 2 of the 22 constants are less than twice as large as their probable errors. Thus, they indicate a real biological relationship between the productions of the half plots in different years. Those which give a higher yield in one year give a higher yield in another year.

For a smaller number of the crops it is possible to divide the original plots into quarter plots, thus securing 184 subplots to be used as a basis of calculation. The coefficients of correlation between the yields in the several years are shown in Table VI.

	1917, oat straw.	1917, total oats.	1918, silage orn.	1919, barley grain.	1919, barley straw.	1919, total barley.
99	−0.116±0.098 −1.18	−0.098±0.098 −1.00	+0.348±0.087 +4.00	−0.539±0.070 −7.66	−0.262±0.092 −2.82	−0.449±0.079 −5.66
05 01	+.166±.097 +1.71	+.229±.094 +2.44	−.071±.099 −0.72	+.527±.071 +7.33	+.341±.087 +3.89	+.483±.076 +6.34
87	+190±.096 +1.98	+.317±.089 +3.56	+.151±.097 +1.56	+.076±.098 +.78	−.003±.099 −.03	+.043±.099 +.43
8m	+.208±.095 +2.19	+.372±.086 +4.33	+.451±.079 +5.71	+.203±.095 +2.13	+.025±.099 +.26	+.131±.097 +1.34
78	+.233±.094 +2.48	+.404±.083 +4.87	+.350±.087 +4.02	+.163±.096 +1.68	+.012±.099 +.13	+.101±.098 +1.03
80	+.281±.092 +3.05	+.429±.081 +4.29	+.209±.095 +2.20	+.255±.092 +2 75	+.139±.097 +1.43	+.221±.094 +2.33
74	+.079±.099 +.80	+.308±.090 +3.42	+.237±.094 +2.52	+.268±.092 +2.90	+.143±.097 +1.47	+.230±.094 +2.44
73	+.188±.096 +1.96	+.395±.084 +4.70	+.242±.094 +2.57	+.283±.091 +3.10	+.153±.097 +1.57	+.244±.093 +2.61
80	+.311±.098 +3.46	+.446±.079 +5.60	+.579±.066 +8.77	+.086±.098 +.87	+.066±.099 +.68	+.084±.098 +.86
71	+.239±.093 +2.55	+.441±.080 +5.50	+.361±.086 +4.18	+246±.093 +2 64	+.139±.097 +1 42	+.215±.094 +2.27
99	+.112±.098 +1.14	+.072±.099 +.73	+.459±.078 +5.88	+.042±.099 +.42	+.184±.096 +1.91	+.119±.098 +1.22
75	+.220±.095 +2.32	+.407±.083 +4.90	+.439±.080 +5.49	+.104±.098 +1.06	+.144±.097 +1.48	+.135±.097 +1 39
	....	....	+.227±.094 +2.41	+.034±.099 +.35	−.020±.099 −.20	+.009±.099 +.10
		....	+.189±.096 +1.97	+.372±.085 +4.34	+.225±.094 +2.38	+.333±.088 +3.76
	....		+.253±.093 +2.72	+.294±.090 +3.24	+.158±.096 +1 63	+.253±.093 +2.71
294	+.189±.096 +1.97	+.253±.093 +2.72		−.166±.096 −1.72	−.063±.099 −.64	−.129±.097 −1.32
299	+.372±.085 +4.34	+.294±.090 +3.24	−.166±.096 −1.72		....	....
299	+.225±.094 +2.38	+.158±.096 +1.63	−.063±.099 −.64	....		....
299	+.333±.088 +3.76	+.253±.093 +2.71	−.129±.097 −1.32	....	....	

TABLE V.—*Interannual correlations for yield of 92 half plots in the Huntley uniform cropping experiment*

	1913, alfalfa I.	1913, alfalfa II.	1913, alfalfa I and II.	1914, alfalfa I.	1914, alfalfa II.	1914, alfalfa I and II.	1915, ear corn.	1916, ear corn.
1913, alfalfa I............				+0.762±0.030 +25.8	+0.543±0.050 +10.9	+0.715±0.034 +20.8	+0.125±0.069 +1.80	+0.686±0.037 +18.4
1913, alfalfa II...........				+.438±.057 +7.71	+.601±.045 +13.4	+.580±.047 +12.4	+.229±.067 +3.43	+.618±.043 +14.2
1913, alfalfa I and II.....				+.707±.035 +20.1	+.673±.038 +17.5	+.763±.029 +25.9	+.208±.067 +3.09	+.768±.029 +26.6
1914, alfalfa I............	+0.762±0.030 +25.8	+.438±.057 +7.71	+.707±.035 +20.1				+.228±.067 +3.42	+.744±.031 +23.7
1914, alfalfa II...........	+.543±.050 +10.9	+.601±.045 +13.4	+.673±.038 +17.5				+.112±.069 +1.61	+.739±.032 +23.2
1914, alfalfa I and II.....	+.715±.034 +20.8	+.580±.047 +12.4	+.763±.029 +25.9				+.185±.068 +2.72	+.820±.023 +35.7
1915, ear corn............	+.125±.069 +1.80	+.229±.067 +3.43	+.208±.067 +3.09	+.228±.067 +3.42	+.112±.069 +1.61	+.185±.068 +2.72		+.339±.063 +5.25
1916, ear corn............	+.686±.037 +18.4	+.618±.043 +14.2	+.768±.029 +26.6	+.744±.031 +23.7	+.739±.032 +23.2	+.820±.023 +35.7	+.339±.063 +5.25	

TABLE VI.—*Interannual correlations for yield of 184 quarter plots in the Huntley uniform cropping experiment*

	1913, alfalfa II.	1914, alfalfa I.	1914, alfalfa II.	1914, alfalfa I and II.	1915, ear corn.	1916, ear corn.
1913, alfalfa II...........		+0.373±0.043 +8.71	+0.373±0.043 +8.71	+0.446±0.040 +11.2	+0.193±0.048 +4.04	+0.512±0.037 +14.0
1914, alfalfa I............	+0.373±0.043 +8.71				+.174±.048 +3.61	+.642±.029 +22.0
1914, alfalfa II...........	+.373±.043 +8.71				+.109±.049 +2.22	+.629±.030 +21.0
1914, alfalfa I and II.....	+.446±.040 +11.2				+.158±.048 +3.27	+.719±.024 +30.0
1915, ear corn............	+.193±.048 +4.04	+.174±.048 +3.61	+.109±.049 +2.22	+.158±.048 +3.27		+.339±.044 +7.42
1916, ear corn............	+.512±.037 +14.0	+.642±.029 +22.0	+.629±.030 +21.0	+.719±.024 +30.0	+.339±.044 +7.42	

Unfortunately the number of crops which can be included in Table VI is rather small. The coefficients are positive in sign throughout, and in all cases they are statistically significant in comparison with their probable errors. The individual constants will receive attention in the following discussion.

The fact that the yields are correlated in the different years for whole plots of 0.17 acre, for half plots of 0.085 acre, and for quarter plots of only 0.0425 acre emphasizes the permanence of the substratum differences. We now have to compare the correlations secured for these three divisions. The difference in the actual magitudes of the correlations appear in Table VII. The three entries, when all comparisons are possible, show: (1) the difference between the correlation for whole plots and half plots, (2) the difference between the correlation for whole plots and quarter plots, and (3) the difference between the correlation for half plots and quarter plots.

The signs are positive when the correlations are larger for the larger areas.

The comparisons show that in general the correlations decrease in magnitude as the areas upon which they are based are subdivided. Thus 16 of the 22 comparisons of the correlations deduced from whole plots and from half plots (first entry) show a lower correlation in the half plots as compared with 6 which show higher correlations in the half plots.

TABLE VII.—*Differences in interannual correlations for whole plots, half plots, and quarter plots*

	1913, alfalfa I.	1913, alfalfa II.	1913, alfalfa I and II.	1914, alfalfa I.	1914, alfalfa II.	1914, alfalfa I and II.	1915, ear corn.	1916, ear corn.
1913, alfalfa I......				−0.0863	−0.0514	−0.0558	+0.0910	+0.0387
1913, alfalfa II........				− .1622	− .1335	− .1462	− .0267	− .1021
				− .2276	− .3615	− .2794	− .0623	− .2081
				− .0653	− .2280	− .1332	− .0355	− .1059
1913, alfalfa I and II..				− .1442	− .1044	− .1152	+ .0403	− .0330
1914, alfalfa I.........	−0.0863	−0.1622	−0.1442				+ .0388	− .0560
		− .2276					− .0150	− .1578
		− .0653					− .0539	− .1018
1914, alfalfa II........	−0.0514	− .1335	− .1044				+ .0705	− .0493
		− .3615					+ .0676	− .1599
		− .2280					− .0029	− .1105
1914, alfalfa I and II.	−0.0558	− .1462	− .1152				+ .0649	− .0377
		− .2794					+ .0385	− .1388
		− .1332					− .0264	− .1011
1915, ear corn.........	+0.0910	− .0267	+ .0403	+ .0388	+ .0705	+ .0649		− .0492
		− .0623		− .0150	+ .0676	+ .0385		+ .0487
		− .0355		− .0539	− .0029	− .0264		− .0004
1916, ear corn.........	+0.0387	− .1021	− .0330	− .0560	− .0493	− .0377	− .0492	
		− .2081		− .1578	− .1599	− .1388	+ .0487	
		− .1059		− .1018	− .1105	− .1011	− .0004	

Of the 12 comparisons possible between the interannual correlations deduced from whole plots and from quarter plots (second entry), 9 show lower correlations for quarter plots as compared with 3 which show higher correlations for the quarter plots. Finally, all 12 of the correlations deduced from quarter plots are lower than the correlations deduced from half plots.

It appears, therefore, that 0.085 and 0.0425 acre are rather too small to give the highest values of the interannual correlations. On areas of this size other factors than the peculiarities of the plots themselves have too large an influence upon variation of yield to allow the individuality of the plots to express itself fully in its influence upon the yields of successive years.

In support of the conclusion that the lower value of the correlations for half and quarter plots is due to the greater variability of the yields of these plots we note that the coefficients of variation for subplots are without exception larger than those for the plots of the original size. The coefficients of variation are as follows for the years in which the plots were subdivided.

Crop.	Whole plots.	Half plots.	Quarter plots.
1913, alfalfa I.	12. 52	14. 93	. . . .
1913, alfalfa II.	13. 60	16. 59	21. 87
1913, alfalfa I and II.	11. 11	13. 34	. . . .
1914, alfalfa I.	17. 94	20. 04	23. 68
1914, alfalfa II.	19. 81	21. 77	25. 87
1914, alfalfa I and II.	17. 47	18. 90	21. 88
1915, ear corn.	7. 29	8. 43	9. 23
1916, ear corn.	13. 43	15. 88	17. 68

It is now desirable to examine the results for the individual crops. In doing this it may be noted that there are two factors to be taken into account. First, there is the possibility of an inherent difference in the plots which is persistent from year to year and is quite independent of the crop grown. Second, it is conceivable that the crop itself may exert an influence upon the soil such that the yields of subsequent crops will be influenced by variations in its growth which are measured in terms of yield.

The first of these factors would influence all correlations between plots—those between the yields of given years and the yields of both preceding and subsequent seasons. The second would influence only correlations with subsequent years.

In a series of only 46 plots it will probably be impossible to distinguish between the influences of these two factors.

We note that the higher yields of beets are followed by lower yields of alfalfa in 1912, but that there is practically no relationship between the yields of sugar beets in 1911 and the yield of other crops on the same

plots from 1913 to 1918. Possible exceptions are ear corn in 1915 and silage corn in 1918, for which the correlations are positive and perhaps statistically significant in comparison with their probable errors. The correlations for yields of sugar beets in 1911 and yields of barley in 1919 are negative in sign and apparently statistically significant in comparison with their probable errors. We have no explanation to offer concerning this apparent relationship. The average value, with regard to sign, of the correlations between the yield of sugar beets and other crops is − 0.077.

The correlations between the 9 different cuttings of alfalfa made during 1912 to 1914 and all other yields are generally positive and statistically significant in comparison with their probable errors. The only exceptions are the negative correlation with sugar beets in 1911 which have already been noted and the slight and statistically insignificant correlation for the 1912 yield of alfalfa and the yield of silage corn in 1918.

Since it is quite reasonable to assume that in a crop harvested more than once a year thickness of stand and variation in the size of the individual plants will have a large influence on the yields of the different plots in the same year, the correlations between the different cuttings of the same year as well as those between single cuttings and totals of two or more cuttings in the same year have been omitted from the tables. The correlations between different cuttings in the same year are given in Table VIII.

TABLE VIII.—*Comparison of correlations between different cuttings of alfalfa in the same year*

Cuttings of alfalfa.	Whole plots.	Half plots.	Quarter plots.
1913, first and second cuttings..	+0. 454±0. 079	+0. 442±0. 057	....
1914, first and second cuttings..	+ .711± .049	+ .633± .042	+0. 558±0. 034
1914, first and third cuttings....	+ .595± .064	....	....
1914 (first plus second) and third cuttings................	+ .653± .057	....	....

We shall now consider the correlations between the yields of alfalfa and between the yields of alfalfa and of other crops on the same plots in different years. The individual constants may be studied in the fundamental table (Table IV). The averages are given in Table IX. This shows that the correlations between different cuttings of alfalfa are on the average larger throughout than those between the yield of alfalfa and the yields of other crops on the same plots.

TABLE IX.—*Comparison of correlations between different yields of alfalfa with correlations between yields of alfalfa and yields of other crops*

Cuttings of alfalfa.	With other cuttings of alfalfa.	With yields of other crops.	Difference.
1912, single cutting...............................	+0. 331	+0. 171	+0. 160
1913, first cutting................................	+ .611	+ .187	+ .424
1913, second cutting..............................	+ .604	+ .282	+ .322
1913, first and second cuttings......................	+ .720	+ .274	+ .446
1914, first cutting................................	+ .666	+ .295	+ .371
1914, second cutting..............................	+ .629	+ .244	+ .385
1914, first and second cuttings......................	+ .699	+ .290	+ .409
1914, third cutting................................	+ .524	+ .303	+ .221
1914, first, second, and third cuttings..............	+ .706	+ .316	+ .390

It is clear, therefore, that either stand or specific adaptation of the individual plots to alfalfa influences to an unusual degree the closeness of correlation between the yields of the plots of alfalfa in the different years.

In the first crop of ear corn (1915) we find higher yields of ear corn in 1916, a negligible difference in the yield of oat grain and straw and total yield in 1917, higher yield of silage corn in 1918, and slightly but not significantly higher yield of barley grain, straw, and total yield in 1919 following higher yield of corn in 1915.

Turning to the constants for ear corn in 1916, we note that higher yields of grain in this year are followed by higher yields of oat straw and grain in 1917 and of silage corn in 1918, and by slightly higher yields of barley grain and straw in 1919.

The average value of the correlation between the yield of ear corn in 1915 and the yield of other crops during the eight years is +0.167, whereas that for ear corn in 1916 and other crops is +0.486. These averages include the correlations for alfalfa, which are, as shown by Table VIII, high for the crop of 1916.

Considering the correlations for oat straw, grain, and total crop on the several plots in 1917 and the yields of silage corn in 1918, we find that higher values of each of these measures of capacity for oat production are on the average followed by slightly, but perhaps not significantly, higher yields of silage corn in 1918 and generally by higher barley yields in 1919.

For the oat yields the average correlations with other crops are: for straw, +0.202; for grain, +0.289; and for total yield, +0.293.

The correlations of the yields of silage corn with the yields of the preceding crops are, with one exception, positive in sign. The average value for the eight years is +0.226.

The averages of the correlations between barley yields and the yields of other crops on the same plots during the eight years of the experiment are +0.141 for grain, +0.086 for straw, and +0.126 for total yield.

Summarizing this discussion of the results for the individual crops, we have the following average values of the correlation coefficients:

1911, sugar beets	−0.077	1915, ear corn	+0.167
1912, total alfalfa	+.242	1916, ear corn	+.486
1913, alfalfa I	+.346	1917, oat straw	+.202
1913, alfalfa II	+.403	1917, oat grain	+.289
1913, alfalfa I and II	+.441	1917, total oats	+.293
1914, alfalfa I	+.401	1918, silage corn	+.226
1914, alfalfa II	+.354	1919, barley grain	+.141
1914, alfalfa I and II	+.407	1919, barley straw	+.086
1914, alfalfa III	+.366	1619, total barley	+.126
1914, alfalfa I to III	+.428	General average	+.274

With the exception of the sugar beets the average correlation for every crop is positive in sign, and in many cases it is of a very material value.

Returning to the averages for the individual crops, we note from Table IX that the lowest correlation for alfalfa, whether with other cuttings of alfalfa or with the yield of other crops, is that for the single cutting of 1912.

It might be suggested that the 1912 yields of alfalfa are less likely to reflect the real producing capacity of the plots than the yields of the later cuttings of this crop, for the reason that the first cutting of alfalfa when sown without a nurse crop is subject to much variation because of slight differences in surface condition of the soil at seeding time and also because of differences in weediness of different plots. Both these conditions would become relatively less important in their effect on crop yield after the first cutting.

Because of its nitrogen-fixing capacity and the resistance to decay of the roots and stubble of alfalfa the correlation between the various yields of this legume and the yields of subsequent crops is of especial interest. Fortunately two crops of ear corn were grown immediately after the alfalfa, which was broken up in the fall of 1914.

A comparison of the correlations of these two series of corn yields with the preceding yields of alfalfa is made in Table X. These coefficients indicate a positive correlation between all the yields of alfalfa and the yields of ear corn in both 1915 and 1916.

Of the 19 correlations determined between the yields of alfalfa for 1912 to 1914 and the yields of ear corn in 1915 only 9 may be looked upon as probably significant in comparison with their probable errors. Of the 19 correlations between the yields of alfalfa in 1912 to 1914 and the yields of ear corn in 1916 only one coefficient—that for the 1912 yield of alfalfa and the 1916 yield of corn—can not be considered as representing a real agronomic relationship between yield of alfalfa and yield of corn.

The constants for 1916 are without exception larger and with two exceptions significantly larger in comparison with their probable errors than those for 1915.

TABLE X.—*Comparison of correlations between yields of alfalfa and yields of the first and of the second subsequent crop of corn*

	Whole plots.			Half plots.			Quarter plots.		
	1915	1916	Difference.	1915	1916	Difference.	1915	1916	Difference.
1912, alfalfa	+0.106±0.098 +1.08	+0.170±0.097 +1.75	+0.064±0.138 +0.46	....	....	....	....	....	....
1913, alfalfa I	+.034±.099 +.34	+.647±.058 +11.2	+.613±.115 +5.34	+0.125±0.069 +1.80	+0.686±.037 +18.4	+0.561±0.078 +7.17	....	....	....
1913, alfalfa II	+.255±.093 +2.74	+.720±.048 +15.0	+.465±.105 +4.45	+.229±.067 +3.43	+.618±.043 +14.2	+.389±.080 +4.89	+0.193±0.048 +4.04	+0.512±0.037 +14.0	+0.319±0.061 +0.53
1913, alfalfa I and II	+.168±.097 +1.73	+.801±.036 +22.3	+.633±.103 +6.12	+.208±.067 +3.09	+.768±.029 +26.6	+.560±.073 +7.67	....	....	....
1914, alfalfa I	+.189±.096 +1.97	+.800±.036 +22.2	+.611±.103 +5.96	+.228±.067 +3.42	+.744±.031 +23.7	+.516±.074 +6.99	+.174±.048 +3.61	+.642±.029 +22.0	+.468±.056 +8.36
1914, alfalfa II	+.041±.099 +.41	+.789±.038 +20.8	+.748±.106 +7.06	+.112±.069 +1.61	+.739±.032 +23.2	+.627±.076 +8.25	+.109±.049 +2.22	+.629±.030 +21.0	+.520±.057 +9.06
1914, alfalfa I and II	+.120±.098 +1.22	+.858±.026 +33.0	+.738±.101 +7.29	+.185±.068 +2.72	+.820±.023 +35.7	+.635±.072 +8.86	+.158±.048 +3.27	+.749±.024 +30.0	+.561±.054 +10.5
1914, alfalfa III	+.329±.088 +3.71	+.710±.049 +14.4	+.381±.101 +3.78	....	....	....	....	....	....
1914, alfalfa I to III	+.191±.095 +2.00	+.880±.022 +39.4	+.689±.098 +7.07	....	....	....	....	....	....
Average	.1592	.7083	.5491	.1811	.7291	.5480	.1585	.6255	.4070

The average value for the nine pairs of correlations deduced from the yields of whole plots is +0.159 for alfalfa and corn yield in 1915 but +0.708 for alfalfa and corn yield in 1916. For the six pairs of correlations which may be deduced for half plots the average of the coefficients for the various yields of alfalfa in 1913 and 1914 and the yield of ear corn in 1915 is +0.181, whereas the average correlation of the same yields of alfalfa with corn one year later is +0.729. Finally, in the four cases in which it was possible to calculate correlations between alfalfa and corn yields on the basis of data for quarter plots the average for the correlations with ear corn in 1915 is +0.159, whereas the constants showing the relationship between the yield of alfalfa in 1913 and 1914 and ear corn in 1916 give an average of +0.626.

This more intimate relationship between the yields of alfalfa and the second crop of ear corn does not necessarily mean that the corn crop of 1916 was larger than that of 1915 but merely that the variations in the individual plot yields in 1916 are more dependent than those of 1915 upon the yields of alfalfa during 1912 to 1914. As a matter of fact the average yield in 1915 was 522.6 pound per plot, while in 1916 it was 396.2 pounds per plot. The greater yield in 1915 may have been, and probably was, due to factors other than soil conditions as such.

It is of interest in this connection to turn back to the table of coefficients of variation of yield (p. 347) and to note that for whole plots, half plots, and quarter plots the coefficients of variation of plot yield are distinctly lower in 1915 than in 1916. This result is quite in line with what one would expect if the fixed nitrogen of the varying growths of alfalfa were not yet fully available in 1915.

There is also another possible explanation for the lower correlation between the alfalfa yields and the yields of corn in 1915. It is always a difficult matter on the heavy soils at Huntley to break up alfalfa sod and to get the soil into good tilth for the succeeding crop. It may be that some of the plots in this field include heavier soil which ordinarily gives good yields but which was harder to get into good condition in time for the 1915 corn crop. If this were the case, these differences in tilth might have been smoothed out by the season's cultivation so as not to be expressed in the 1916 crop yields.

Some light may be thrown upon the problem of the residual influence of alfalfa in the following manner.

If the correlations between the plot yields of later crops be in a large degree determined by differences in fertility referable to differences in stand and yield of the preceding alfalfa crops, one might expect a closer correlation between the yields of ear corn in 1916 and oats in 1917 than between ear corn in 1916 and ear corn in 1915, since, as is shown above, variations in the alfalfa yields have little influence until 1916. This will be true, provided there be a residual influence of the variations in the yields of alfalfa such that these variations in fertility due to varia-

tions in yield from 1912 to 1914 inclusive will influence not merely the yield of corn in 1916 but the yield of oats in 1917, etc. The correlations between corn yields in 1915 and corn yields in 1916 and the yields of subsequent crops are shown side by side in Table XI.

TABLE XI.—*Comparison of correlations of the yields of ear corn in 1915 and in 1916 with the yields of subsequent crops*

	Corn 1915.	Corn, 1916.	Difference.
1917.			
Oat grain................	−0.025±0.099	+0.497±0.075	+0.522±0.124
Oat straw................	+.112±.098	+.220±.095	+.108±.136
Total yield...............	+.072±.099	+.407±.083	+.335±.129
1918.			
Silage corn...............	+.459±.078	+.439±.080	−.020±.112
1919.			
Barley grain..............	+.042±.099	+.104±.098	+.062±.139
Barley straw.............	+.184±.096	+.144±.097	−.040±.136
Total yield...............	+.119±.098	+.135±.097	+.016±.138

These comparisons show that the yields of oats in 1917 are much more closely correlated with the yields of corn in 1916 than with the yields of ear corn in 1915. No such relationship is apparent in the correlations for silage corn in 1918 or for barley in 1919. The after effect of the alfalfa crops of 1912 to 1914 is, therefore, apparently largely limited to an influence on the yield of oats in 1917.

Turning from this indirect to a more direct method of comparison, we have determined the averages of the correlations between the several individual cuttings of alfalfa and the yields of the single antecedent and of the five subsequent crops. The results are given in Table XII.

TABLE XII.—*Averages of the correlations between the cuttings of alfalfa in 1912 to 1914 and the antecedent and succeeding crops*

Crop correlated with alfalfa.	Grain.	Straw.	Total yield.
Sugar beets, 1911...................	....	....	−0.082
.	.	.	.
.	.	.	.
.	.	.	.
Ear corn, 1915....................	+0.159	....	....
Ear corn, 1916....................	+.708	....	....
Oats, 1917......................	+.437	+0.210	+.371
Silage corn, 1918.................	....	....	+.279
Barley, 1919.....................	+.234	+.113	+.195

There should be no correlation between the yield of sugar beets and alfalfa except that due to the initial heterogeneity of the field. The

insignificant negative correlation observed may be due to some pecu-
liarity of the crop. The comparison of the correlation for the 1915 and
1916 corn crops has already been made (Table XI). Inspection of the
averages in Table XII shows that on whatever character they are based
the correlations decrease from the maximum relationship observed in
1916 to the lowest values in 1919.

Whether the residual influence of alfalfa *per se* has any influence on
the 1919 or later crops can only be determined by further experimenta-
tion in which the interannual correlations can be deduced from the
yields of plots upon which alfalfa has not been grown.

III.—DISCUSSION AND RECAPITULATION

The purpose of this paper has been to present the results of a new
method of attack upon the problems of (a) the permanency of the differ-
ences which are found in the plots of an experimental field, and of (b)
the influence of variations in the yields of certain crops in the rotation
upon the yields of subsequent crops.

The data upon which the studies were primarily based comprise the
yields of 46 plots—subdivided in several cases into half plots and quarter
plots—each of o.17 acre in area at the Huntley (Mont.) Field Station of
the Office of Western Irrigation Agriculture for the nine years between
1911 and 1919, inclusive.

The uniform cropping experiment, involving sugar beets, alfalfa,
corn, oats, and barley, was initiated merely to determine the variation
in the yields of plots of a given size when homogeneously planted and
uniformly treated. The experimental procedure was, therefore, deter-
mined in advance and was wholly independent of the statistical analysis.
This is in certain regards fortunate. It frees the data absolutely from
any suspicion of an influence of preconceptions or of personal equation
on the biometric results. On the other hand, it is quite possible after
the statistical analyses have been made to recognize ways in which the
experiments could have been improved and made to yield more valuable
results. This is, however, a feature of research in general. The dis-
covery of inadequacies in a first set of experiments makes possible their
elimination in subsequent work. The most unfortunate defect in the
data was that the harvesting and weighing could not be done by half
and quarter plots in 1917, 1918, and 1919, but this curtailment could
not be avoided under existing conditions.

The results of a previous study (3) have shown that fields selected for
plot tests of all kinds are practically without exception heterogeneous
to a degree that influences profoundly the yields of the crops grown
upon them. It was there pointed out that the correlation between
the yields of adjacent plots might either be due to initial physical and
chemical differences in the soil or be referable to the influence of previ-
ous crops upon the composition, texture, or tilth of the soil.

The first purpose of the present study has been to determine whether such differences in fields selected for their apparent uniformity by skilled agronomists are of a purely transitory nature or whether they are of a relatively permanent character.

This problem can be solved by determining whether in such series of uniformly treated plots the yields of the same plots in different years are correlated.

The results of the present study show that of the 152 correlations between the yields of the plots in different years, 133 are positive as compared with 19 which are negative in sign. The average value of the positive correlations is + 0.335, whereas the average of the negative constants is − 0.148. The general average is + 0.274. With the exception of the 1911 crop of sugar beets the correlation between the yields of each individual crop and the yields on the same plots in the eight other years of the experiment are on the average positive.

The data available for half and quarter plots fully substantiate the results for whole plots.

The results show conclusively, therefore, that plots, even of the small size and apparent uniformity of those at the Huntley Station, are characterized by differences which may persist throughout a period of years. Thus, in general, plots which produce more in one year will produce more in another year.

This is, of course, a well-recognized principle for large tracts. Its validity for small plots has apparently not been recognized heretofore. It is probably not a principle of universal applicability, because of the fact that meteorological as well as soil conditions play a large part in determining yield. It is quite probable that certain soil characteristics would result in maximum yields with one set of meteorological conditions but in minimum yields with another complex of aereal conditions.

The determination of the proximate factors to which these correlations are due presents a problem of considerable difficulty. Unfortunately (for this phase of the problem only) alfalfa was introduced early in the rotation and occupied the ground for three of the nine years covered by the experiment. It seems quite possible that the correlations between certain of the yields is due in part to the variation in nitrogen content of the soil referable to the variation in thickness of stand and strength of growth of the alfalfa crops.

The results show that there is but little correlation between the alfalfa yields of 1912 to 1914 and the ear corn yields of 1915, whereas the correlations for ear corn in 1916 are high. Thus the influence of alfalfa upon the yield of a subsequent crop is not fully evident until the second year after it is turned under.

There is a definitely demonstrable residual influence of the variation of alfalfa yields upon the yields of subsequent crops. The influence of

the alfalfa upon the yield of subsequent crops decreases with the lapse of time from the maximum correlation found for ear corn in 1916. The residual influence of the alfalfa is clearly marked in the oat crop of 1917 and may still be evident in the silage corn and barley crops of 1918 and 1919.

In view of the early introduction of alfalfa into the rotation, it is impossible to determine whether the correlations between yields other than those for alfalfa are due to the variation from plot to plot of the amount of nitrogen fixed by the alfalfa or whether it is to a considerable extent due to the original heterogeneity of the plots. This and other problems which will suggest themselves to the reader can be solved only by the analysis of further experimental data. The illustrations of the present paper are sufficient to show the value of the application of the interannual correlation method to agronomic problems.

LITERATURE CITED

(1) HARRIS, J. Arthur.
 1915. ON A CRITERION OF SUBSTRATUM HOMOGENEITY (OR HETEROGENEITY) IN FIELD EXPERIMENTS. *In* Amer. Nat., v. 49, no. 583, p. 430–454.
(2) ————
 1915. THE VALUE OF INTER-ANNUAL CORRELATIONS. *In* Amer. Nat., v. 49, no. 587, p. 707–712.
(3) ————
 1920. PRACTICAL UNIVERSALITY OF FIELD HETEROGENEITY AS A FACTOR INFLUENCING PLOT YIELDS. *In* Jour. Agr. Research, v. 19, no. 7, p. 279–314. Literature cited, p. 313–314.
(4) LEHMANN, A.
 1907. SEVENTH ANNUAL REPORT OF THE AGRICULTURAL CHEMIST FOR THE YEAR 1905–1906. [Department of Agriculture, Mysore State.] 53 p. Bangalore.
(5) LYON, T. L.
 1912. SOME EXPERIMENTS TO ESTIMATE ERRORS IN FIELD PLAT TESTS. *In* Proc. Amer. Soc. Agron., v. 3, p. 89–114, 5 fig.
(6) SMITH, Louie H.
 1910. PLOT ARRANGEMENTS FOR VARIETY EXPERIMENTS WITH CORN. *In* Proc. Amer. Soc. Agron., v. 1, 1907/09, p. 84–89.
(7) STOCKBERGER, W. W.
 1912. A STUDY OF INDIVIDUAL PERFORMANCE IN HOPS. *In* Ann. Rpt. Amer. Breeders' Assoc., v. 7/8, p. 452–457.
(8) ————
 1916. RELATIVE PRECISION OF FORMULAE FOR CALCULATING NORMAL PLOT YIELDS. *In* Jour. Amer. Soc. Agron., v. 8, no. 3, p. 167–175.

TISSUE WEIGHT AND WATER CONTENT IN A TETRACOTYLEDONOUS MUTANT OF PHASEOLUS VULGARIS.

BY J. ARTHUR HARRIS.

105 (1687)

Tissue weight and water content in a tetracotyledonous mutant of Phaseolus vulgaris.

By J. ARTHUR HARRIS.

[*From the Station for Experimental Evolution, Cold Spring Harbor, L. I.*]

In 1915 studies of the tissues of normal and variant bean seedlings were begun in an effort to explain the differential mortality with respect to morphological characters demonstrated in an earlier experiment.[1]

In a first paper[2] it was demonstrated that teratological seedlings show a lower capacity for the development of primordial leaf tissue than do normal individuals of the same strain grown under conditions as nearly as possible identical. In these first experiments the conclusions were based upon the green weight of primordial leaves.

In a second study, tissue weight determinations were based on the trifoliate leaves of the third node as well as on the primordial leaves of the second node.[3] In these two investigations we attempted to determine to what extent morphologically aberrant seedlings differ from the normal seedlings of the race to which they belong in their physiological characters in so far as these can be measured by the capacity for the production of tissue. The results indicated that teratological seedlings show a lower capacity

[1] Harris, J. Arthur, *Science*, 1912, N. S., xxxvi, 713–715.

[2] Harris, J. Arthur, *Genetics*, 1916, i, 185–196.

[3] Harris, J. Arthur, Brooklyn Botanic Garden Memoirs, 1918, i, 167–174.

for tissue production as measured both by green weight and dry weight in primordial and first compound leaves than do their normal controls.

In a subsequent series of investigations we have instituted comparisons between the highly abnormal seedlings of a tetra-cotyledonous race of Phaseolus[1] and the normal seedlings of the parental race from which it originated.

The tetracotyledonous race is characterized by a modal number of four cotyledons and four primordial leaves but both of these characters are highly variable.

Classifying the tetracotyledonous plants according to number of primordial leaves, we have the mean green weight and the mean dry weight of primordial leaf tissue in teratological and normal seedlings shown in the accompanying table.

The data are given as average weights per plant and per leaf. The average per cent. of dry substance is shown in the final column of the table.   All the values are averages of constants based on samples of approximately 100 plants.

The data show that without exception the mean green weight and the mean dry weight per plant of primordial leaf tissue is lower in the tetracotyledonous race than in the normal race.

The mean percentage differences (obtained by using the constants for the normal plants as a base) for green weight per plant range from -3.10 for the plants with six primordial leaves[2] to -31.55 per cent. for group of plants with 2 primordial leaves.

The percentage differences for dry weight of primordial leaves in tetracotyledonous and dicotyledonous races vary from -7.93 per cent. for the group of plants with 6 primordial leaves to -32.55 for plants of the tetracotyledonous race with 2 primordial leaves.

It will be noted that the difference between the abnormal and the normal plants decreases as the number of primordial leaves on the abnormal plants increases.

The results for the average green and dry weight per leaf in the mutant and normal series fully substantiate the conclusions

---

[1] Harris, J. Arthur, *Proceedings National Academy of Science*, 1916, ii, 317–318.
    Harris, J. Arthur, Memoir, N. Y. Botanical Garden, 1916, vi, 229–244.
[2] Theoretically plants with 7 leaves should have shown a smaller difference than seedlings with six leaves but the number of seedlings available was not so large and the constant is therefore not as trustworthy.

concerning the physiological differentiation of the two races to be drawn from the average weights per plant.

The differences in the average percentage of dry substance vary considerably but it is impossible to state in the absence of probable errors that the ratio differs from class to class.

The foregoing results for a heritable abnormal race substantiate the conclusions concerning the association of physiological with morphological differences already drawn from a comparison of variant and normal individuals within the same race.

Number of Leaves of Abnormal Race.	Pairs of Plants.	Value per Plant.		Value per Leaf.		Per cent. Dry Substance.
		Mean Green Weight.	Mean Dry Weight.	Mean Green Weight.	Mean Dry Weight.	
2 Leaves						
Abnormal............	196	.52390	.03790	.26195	.01900	7.3090
Control.............	196	.76285	.05645	.38145	.02825	7.4350
Difference...........		−.23895	−.01855	−.11950	−.00925	−.1260
Percentage difference...		−31.55	−32.55	−31.55	−32.40	
3 Leaves						
Abnormal............	500	.61086	.04174	.20360	.01392	6.8516
Control.............	500	.77094	.05606	.38550	.02806	7.2842
Difference...........		−.16008	−.01432	−.18190	−.01414	−.4326
Percentage difference...		−20.74	−25.54	−47.16	−50.38	
4 Leaves						
Abnormal............	700	.68947	.04587	.17238	.01147	6.6918
Control.............	700	.78631	.05501	.39318	.02751	7.0305
Difference...........		−.09684	−.00914	−.22080	−.01604	−.3387
Percentage difference...		−11.97	−16.34	−55.95	−58.12	
5 Leaves						
Abnormal............	500	.70542	.04740	.14108	.00948	6.7618
Control.............	500	.77152	.05504	.38580	.02754	7.1692
Difference...........		−.06610	−.00764	−.24472	−.01806	−.4074
Percentage difference...		−8.60	−13.78	−63.46	−65.52	
6 Leaves						
Abnormal............	300	.74746	.05146	.12456	.00860	6.9076
Control.............	300	.77206	.05596	.38603	.02796	7.2616
Difference...........		−.02460	−.00446	−.26146	−.01936	−.3540
Percentage difference...		−3.10	−7.93	−67.66	−69.16	
7 Leaves						
Abnormal............	138	.73120	.05570	.10450	.00800	7.6200
Control.............	138	.77850	.06140	.38930	.03070	7.8920
Difference...........		−.04730	−.00570	−.28480	−.02270	−.2720
Percentage difference...		−6.00	−16.76	−73.10	−73.90	

LEAF-TISSUE PRODUCTION AND WATER CONTENT IN A MUTANT RACE OF PHASEOLUS VULGARIS

J. ARTHUR HARRIS

LEAF-TISSUE PRODUCTION AND WATER CONTENT IN A MUTANT RACE OF PHASEOLUS VULGARIS

J. ARTHUR HARRIS

Introductory

In a preceding paper[1] it was shown that the survival of the bean plant is in a measurable degree dependent upon the morphological characteristics of the seedling. In 1915 a series of investigations was undertaken to determine, if possible, something of the proximate causes of the differential death rate. It was also hoped that some light would thereby be thrown upon the proximate causes underlying the occurrence of teratological variations in the seedlings of *Phaseolus*. In undertaking this work the assumption seemed justified that if innate physiological conditions which might affect growth be associated with morphological variations, some influence of these factors should be recorded in the size or other characteristics which result from the relatively enormous expansion which the organs of the embryo undergo in the course of germination and the establishment of the seedling.

A first study[2] demonstrated that teratological seedlings in general show a lower capacity for the development of primordial leaf tissue than do normal seedlings grown under as nearly as possible identical conditions. The data then available indicated that a reduction of the volume of primordial leaf tissue is associated with abnormalities of all the abnormal types studied, but that the type of variation influences in some degree the amount of reduction. In these first experiments the conclusions were based on primordial leaves only. The use of such leaves has the obvious disadvantage that they are formed in the seed, and undergo merely an enormous expansion (and possibly a little differentiation) in the germination

[1] HARRIS, J. ARTHUR, A simple demonstration of the action of natural selection. Science, N.S. **36**:713–715. 1912.

[2] ———, Studies on the correlation of morphological and physiological characters: The development of the primordial leaves in teratological bean seedlings. Genetics **1**:185–196. 1916.

of the seed and the development of the plantlet to the stage at which measurements were made. Since the development of the primordial leaves during the germination and establishment of the seedling is relatively great, it seemed quite legitimate to use the weight of green tissue produced by these leaves as a measure of the physiological capacity of seedlings of various types. The fact that these leaves are differentiated in the seed, however, constitutes a valid objection against their use as a sole measure of the physiological capacity of the seedling. For such purposes a constant based upon some organ developed later seemed desirable.

In a second study,[3] therefore, the tissue weight determinations were extended to the trifoliate leaves of the third node, as well as to the primordial leaves of the second node. This leaf was used because groups of plants of more uniform development can be selected at the time of maturity of this leaf, than at any later stage, and because the first compound leaf reaches a degree of maturity sufficient for the purpose of the present study before the primordial leaves are too old to be used. It is possible, therefore, to check results by determinations made on organs differentiated both in age and in structure. In the first investigation the green weight of the leaf tissue served as the fundamental measurement. In addition to this character certain measurements on the sap properties were also made. In the study of the saps some difficulties were encountered, and it seemed desirable to discontinue that phase of the work temporarily and to carry out determinations of dry weight and water content instead. The present study, therefore, has to do only with the green weight, the dry weight, and the percentage of dry matter.

Recent investigations fell into two phases. The first was an endeavor to determine to what extent seedlings which are morphologically aberrant in the race to which they belong also differ from the normal seedling of the race in their physiological characters, in so far as these can be measured by the capacity for the production of tissue. In the second the investigation was extended from intra-racial to inter-racial comparisons, to ascertain if possible to what

[3] HARRIS, J. ARTHUR, Further studies on the interrelationship of morphological and physiological characters in seedlings of *Phaseolus*. Brooklyn Bot. Gard. Mem. 1:167–174. 1918.

extent a highly abnormal race differs from the parental strain from which it originated.

Materials and methods

In this paper the characteristics of a fully heritable teratological race are considered. The material was furnished by a tetracotyledonous race, the origin and general characteristics of which have been considered elsewhere.[4] The tissue of plants of the tetracotyledonous race were compared with those of the normal line from which it originated.

Seeds of the two series grown in the same field in 1917 were germinated in flats of sand in 1919. Four lots of fifty seeds each, two of the tetracotyledonous plants and two of the normal ancestral line, were germinated in alternate positions in the same flat. Conditions, therefore, were as nearly comparable as possible in the germination of the two series. When the seedlings were of the proper size for potting, one seedling of the tetracotyledonous race and one normal control taken from the same flat were transferred to 3-inch pots of soil, where they stood until they were ready for the collection of samples of tissues. Weighings were then made of the primordial leaves in the two cases. Thus, although weight and other characteristics vary from sample to sample because of age and the innumerable slight influences of significance in growth, the aberrant plants and their controls from the very beginning had as nearly as possible identical environment. However much the pairs combined in the same sample may differ among themselves, there seems no possibility of considering that the differentiation here shown to exist between the morphologically typical and the morphologically aberrant individuals is due to any extrinsic cause. In the absence of any knowledge of the amount of variation in the characteristics of the leaves to be investigated, it was impossible to compute in advance the size of the sample which should be taken. Accordingly it was arbitrarily fixed as 100 plants.[5]

[4] HARRIS, J. ARTHUR, A tetracotyledonous race of *Phaseolus vulgaris*. Mem. N.Y. Bot. Gard. **6**:229–244. 1916.

———, De Vriesian mutation in the garden bean *Phaseolus vulgaris*. Nat. Acad. Sci. **2**:317–318. 1916.

[5] Sample 305 contained 96 plants, sample 252 only 81 plants, and sample 259 contained 138 plants.

In work with the variants in normal lines of beans there is no difficulty whatever in distinguishing primordial leaves from those subsequently formed, except occasionally in extreme variations involving stem characters such as would ordinarily be classed as fasciations. In the case of the tetracotyledonous race, however, it is often difficult to distinguish between true primordials (those formed in the seed) and the simple leaves (not compound) formed subsequently. This difficulty was noted in the first paper on the tetracotyledonous race, and two series of countings at different stages of development of the seedling were made to determine to what extent personal equation may affect the constants for number of primordial leaves.

For practical reasons it was not feasible to count the leaves of the tetracotyledonous plants used in these experiments immediately after germination. Countings, therefore, were made just before the samples were taken. The numbers recorded are those of leaves which were regarded as certainly primordial. Those which from their color or texture appeared to be of subsequent development were omitted. In this race filaments are of rather frequent occurrence. These are probably morphologically much reduced leaves, and were also disregarded. Thus the number of primordial leaves is probably on the average slightly under rather than over the true number for the series as a whole. Since we are primarily concerned with a comparison between definite types of seedling classification with respect to number of leaves, this procedure can introduce no sensible error into the results.

Because of some uncertainty as to the leaves which were to be considered primordial and the considerable variation in the stages of development of the compound leaves in the tetracotyledonous plants, it did not seem feasible in the majority of determinations to consider separately the weight of tissue formed by the compound leaves. This, however, has been done indirectly in the case of certain samples based on the plants as a whole.

Data

The data fall in three groups: a series of weighings of primordial leaves of plants unclassified with respect to number of primordial leaves; a series of weighings of primordial leaves of plants classified

with respect to number of primordial leaves; and a series of weighings of total epicotyledonary tissue.

TABLE I

MEAN GREEN WEIGHT PER PLANT AND PER LEAF IN SEEDLINGS OF A TETRACOTYLEDONOUS RACE AND IN NORMAL PLANTS OF THE ANCESTRAL RACE

SAMPLE	VALUES PER PLANT				VALUES PER LEAF			
	Abnormal	Control	Difference	Percentage difference	Abnormal	Control	Difference	Percentage difference
Unclassified								
226......	0.6991	0.7516	−0.0525	− 6.9	0.1718	0.3758	−0.2040	−54.2
227......	0.6972	0.7607	−0.0635	− 8.3	0.1680	0.3804	−0.2124	−55.8
228......	0.6323	0.7568	−0.1245	−16.4	0.1542	0.3784	−0.2242	−59.2
229......	0.7012	0.6862	+0.0150	+ 2.1	0.1793	0.3431	−0.1638	−47.7
2 leaves								
258......	0.4520	0.7263	−0.2743	−37.7	0.2260	0.3632	−0.1372	−37.7
305......	0.5958	0.7994	−0.2036	−25.4	0.2979	0.3997	−0.1018	−25.4
3 leaves								
255......	0.6313	0.7760	−0.1447	−18.6	0.2104	0.3880	−0.1776	−45.7
273......	0.5662	0.7189	+0.1527	−21.2	0.1887	0.3595	−0.1708	−47.5
283......	0.5791	0.7766	−0.1975	−25.4	0.1930	0.3883	−0.1953	−50.2
303......	0.6577	0.8015	−0.1438	−17.9	0.2192	0.4008	−0.1816	−45.3
319......	0.6200	0.7817	−0.1617	−20.6	0.2067	0.3909	−0.1842	−47.1
4 leaves								
253......	0.6703	0.7786	−0.1083	−13.9	0.1676	0.3893	−0.2217	−56.9
264......	0.5994	0.6671	−0.0677	−10.1	0.1499	0.3336	−0.1837	−55.0
286......	0.7066	0.8103	−0.1037	−12.7	0.1767	0.4052	−0.2285	−56.3
296......	0.6556	0.8142	−0.1586	−19.4	0.1639	0.4071	−0.2432	−59.7
298......	0.7368	0.9028	−0.1660	−18.3	0.1842	0.4514	−0.2672	−59.1
314......	0.7201	0.7783	−0.0582	− 7.4	0.1800	0.3892	−0.2092	−53.7
320......	0.7375	0.7529	−0.0154	− 2.0	0.1844	0.3765	−0.1921	−51.0
5 leaves								
254......	0.6371	0.7639	−0.0998	−13.5	0.1274	0.3685	−0.2411	−65.4
268......	0.6409	0.6987	−0.0578	− 8.2	0.1282	0.3494	−0.2212	−63.3
287......	0.7334	0.7867	−0.0533	− 6.7	0.1467	0.3934	−0.2467	−62.7
300......	0.8032	0.8366	−0.0334	− 3.9	0.1606	0.4183	−0.2577	−61.6
321......	0.7125	0.7987	−0.0862	−10.7	0.1425	0.3994	−0.2569	−64.3
6 leaves								
256......	0.7123	0.7268	−0.0145	− 1.9	0.1187	0.3634	−0.2447	−67.3
289......	0.7351	0.7646	−0.0295	− 3.8	0.1225	0.3823	−0.2598	−67.9
307......	0.7950	0.8248	−0.0298	− 3.6	0.1325	0.4124	−0.2799	−67.8
7 leaves								
259......	0.7312	0.7785	−0.0473	− 6.0	0.1045	0.3893	−0.2848	−73.1
Epicotyl								
244......	1.4388	2.0369	−0.5981	−29.4	......	......	......	......
246......	1.5442	1.9133	−0.3691	−19.3	......	......	......	......
248......	1.5448	2.0396	−0.4948	−24.3	......	......	......	......
250......	1.6325	1.9028	−0.2703	−14.2	......	......	......	......
252......	1.0634	1.1512	−0.0878	− 7.6	......	......	......	......

PLANTS UNCLASSIFIED WITH RESPECT TO NUMBER OF PRIMORDIAL LEAVES.—In preliminary work (samples 226–229) the total weight of primordial leaf tissue in the abnormal seedlings is compared

with the total weight in the control plants irrespective of the number of primordial leaves formed by the individual plants of the tetracotyledonous race. The total number of leaves per plant, however, was determined in these four series.[6] Thus it is possible to give the average weights both per plant and per leaf in the two series. The results show that in three of the four cases the green weight as given in table I of the approximately four primordial leaves of the tetracotyledonous race is lower than that of the two primordial leaves of the dicotyledonous strain. The percentage differences in total weight range from $+2.1$ to -16.4, with a general average of -7.37. When the comparison is made on the basis of mean weight per leaf, the primordial leaf of the abnormal seedling is found to be on the average 54.22 per cent lighter than the leaf of the normal seedling.

For dry weight, given in table II, all four series show lower average weight in the tetracotyledonous strain. The percentage differences for dry weight of primordial leaves per plant vary from -1.6 to -18.0, with a general average of -10.90. On the basis of mean dry weight per leaf, the weight for tetracotyledonous plants is found to be from 49.6 to 59.9 per cent lower than that of the normal seedling, with a general average percentage difference of -55.92. Thus the results for these four samples clearly indicate that an abnormal race shows the same relationship to the normal parental race as do abnormal individual seedlings to the normal seedlings in the same race.

PLANTS CLASSIFIED WITH RESPECT TO NUMBER OF PRIMORDIAL LEAVES.—Upon the completion of this preliminary comparison it seemed worth while to analyze the relationships more minutely by considering individually the results for seedlings of the tetra-cotyledonous race with varying numbers of primordial leaves. These results were only attained at the cost of great labor, since it was difficult to secure considerable numbers of seedlings of any given type simultaneously. It was necessary, therefore, to make determinations for abnormal and control plants in small sub-samples, and to combine these to form samples of 100 seedlings

---

[6] The average numbers per plant were as follows in the four samples: $226 = 4.07$, $227 = 4.15$, $228 = 4.10$, and $229 = 3.91$.

each. The results are shown in table I for green weight, table II for dry weight, and in table III for the percentage of dry matter in the primordial leaves. The data show that in the case of both green and dry weight tissue production is invariably higher in the

TABLE II

MEAN DRY WEIGHT PER PLANT AND PER LEAF IN SEEDLINGS OF A TETRACOTYLEDONOUS RACE AND IN NORMAL PLANTS OF THE ANCESTRAL RACE

SAMPLE	VALUES PER PLANT				VALUES PER LEAF			
	Abnormal	Control	Difference	Percentage difference	Abnormal	Control	Difference	Percentage difference
Unclassified								
226......	0.0529	0.0600	−0.0071	−11.8	0.0130	0.0300	−0.0170	−56.6
227......	0.0530	0.0604	−0.0074	−12.2	0.0128	0.0302	−0.0174	−57.6
228......	0.0528	0.0644	−0.0116	−18.0	0.0129	0.0322	−0.0193	−59.9
229......	0.0539	0.0548	−0.0009	− 1.6	0.0138	0.0274	−0.0136	−49.6
2 leaves								
258......	0.0355	0.0587	−0.0232	−39.5	0.0178	0.0294	−0.0116	−39.4
305......	0.0403	0.0542	−0.0139	−25.6	0.0202	0.0271	−0.0069	−25.4
3 leaves								
255......	0.0492	0.0649	−0.0157	−24.1	0.0164	0.0325	−0.0161	−49.5
273......	0.0427	0.0562	−0.0135	−24.0	0.0142	0.0281	−0.0139	−49.4
283......	0.0404	0.0565	−0.0161	−28.4	0.0135	0.0283	−0.0148	−52.2
303......	0.0404	0.0539	−0.0135	−25.0	0.0135	0.0270	−0.0135	−50.0
319......	0.0360	0.0488	−0.0128	−26.2	0.0120	0.0244	−0.0124	−50.8
4 leaves								
253......	0.0517	0.0641	−0.0124	−19.3	0.0129	0.0321	−0.0192	−59.8
264......	0.0470	0.0532	−0.0062	−11.6	0.0118	0.0266	−0.0148	−55.6
286......	0.0493	0.0576	−0.0083	−14.4	0.0123	0.0288	−0.0165	−57.2
296......	0.0409	0.0540	−0.0131	−24.2	0.0102	0.0270	−0.0168	−62.2
298......	0.0444	0.0568	−0.0124	−21.8	0.0111	0.0284	−0.0173	−60.9
314......	0.0444	0.0516	−0.0072	−13.9	0.0111	0.0258	−0.0147	−56.9
320......	0.0434	0.0478	−0.0044	− 9.2	0.0109	0.0239	−0.0130	−54.3
5 leaves								
254......	0.0488	0.0604	−0.0116	−19.2	0.0098	0.0302	−0.0204	−67.5
268......	0.0462	0.0537	−0.0075	−13.9	0.0092	0.0269	−0.0177	−65.7
287......	0.0501	0.0562	−0.0061	−10.8	0.0100	0.0281	−0.0181	−64.4
300......	0.0496	0.0549	−0.0053	− 9.6	0.0099	0.0275	−0.0176	−64.0
321......	0.0423	0.0500	−0.0077	−15.4	0.0085	0.0250	−0.0165	−66.0
6 leaves								
256......	0.0528	0.0558	−0.0030	− 5.3	0.0088	0.0279	−0.0191	−68.4
289......	0.0520	0.0555	−0.0035	− 6.3	0.0087	0.0277	−0.0190	−68.5
307......	0.0496	0.0565	−0.0069	−12.2	0.0083	0.0283	−0.0200	−70.6
7 leaves								
259......	0.0557	0.0614	−0.0057	− 9.2	0.0080	0.0307	−0.0227	−73.9
Epicotyl								
244......	0.1048	0.1507	−0.0459	....	....	....	....	....
246......	0.1159	0.1443	−0.0284	....	....	....	....	....
248......	0.1127	0.1514	−0.0387	....	....	....	....	....
250......	0.1193	0.1413	−0.0220	....	....	....	....	....
252......	0.0939	0.1045	−0.0106	....	....	....	....	....

two primordial leaves of the normal ancestral strain than it is in the two to seven leaves of the tetracotyledonous strain. The percentage

TABLE III

PERCENTAGE DRY SUBSTANCE IN SEEDLINGS OF A TETRACOTYLE-
DONOUS RACE AND IN NORMAL PLANTS OF THE
ANCESTRAL RACE

SAMPLE	PRIMORDIAL LEAVES		
	Abnormal	Control	Difference
Unclassified			
226.............	7.566	7.982	−0.416
227.............	7.601	7.940	−0.339
228.............	8.350	8.509	−0.159
229.............	7.686	7.986	−0.300
2 leaves			
258.............	7.853	8.082	−0.229
305.............	6.765	6.788	−0.023
3 leaves			
255.............	7.793	8.363	−0.570
273.............	7.541	7.817	−0.276
283.............	6.976	7.275	−0.299
303.............	6.142	6.724	−0.582
319.............	5.806	6.242	−0.436
4 leaves			
253.............	7.712	8.232	−0.520
264.............	7.841	7.974	−0.133
286.............	6.977	7.108	−0.131
296.............	6.238	6.632	−0.394
298.............	6.026	6.291	−0.265
314.............	6.165	6.629	−0.464
320.............	5.884	6.348	−0.464
5 leaves			
254.............	7.659	8.196	−0.537
268.............	7.208	7.685	−0.477
287.............	6.831	7.143	−0.312
300.............	6.175	6.562	−0.387
321.............	5.936	6.260	−0.324
6 leaves			
256.............	7.412	7.677	−0.265
389.............	7.073	7.258	−0.185
307.............	6.238	6.850	−0.612
7 leaves			
259.............	7.620	7.892	−0.272
Epicotyl			
244.............	.7.283	7.398	−0.115
246.............	7.505	7.541	−0.036
248.............	7.295	7.423	−0.128
250.............	7.307	7.425	−0.118
252.............	8.834	9.083	−0.249

values show considerable variation from sample to sample. As might have been expected on a priori grounds, the deficiency of

he weight of primordial leaves in the tetracotyledonous line is greatest when only two leaves are formed.

A comparison of the average percentage differences for abnormal plants with various numbers of leaves gives the following results:

No. of leaves	Green weight	Dry weight
2	−31.55	−32.55
3	−20.74	−25.54
4	−11.97	−16.34
5	− 8.60	−13.78
6	− 3.10	− 7.93

Only one sample is available for seedlings with seven primordial leaves, and it is omitted from the comparison. The results for the other five classes show that:

a) The difference between the total weight of primordial leaf tissue in the abnormal seedling and its normal control decreases as the number of leaves in the abnormal plant increases, but that throughout the entire range of variation of leaf number studied the tetracotyledonous plant produces a smaller total weight of leaf tissue than do normal plants of the line from which it was derived.

b) The differences between tetracotyledonous and dicotyledonous plants are always greater when the comparison is made on the basis of dry weight than when it is made on the basis of green weight.

If the comparison be made on the basis of average weight per leaf the following results are obtained:

No. of leaves	Green weight	Dry weight
2	−31.55	−32.40
3	−47.16	−50.38
4	−55.95	−58.12
5	−63.46	−65.52
6	−67.66	−69.16

The percentage differences in average weight per leaf of course increase as the number of leaves in the abnormal seedlings increases. Again the greater percentage difference when dry weight serves as a basis of comparison is conspicuous. The percentage of dry matter

in the seedlings of this extremely abnormal race is shown in comparison with the normal control plants in table III. The results are self-explanatory. Without exception, in the twenty-three samples representing weighings of 9958 leaves of abnormal and 4668 leaves of normal plants, the percentage of dry matter is lower in the abnormal than in the control series. A study of the averages for the individual groups of seedlings, classified with respect to primordial leaf number, does not suggest a significant difference in the percentage of dry matter in the different classes of seedlings. Probably a far larger series of weighings would be required to bring out such a differentiation if it exists at all.

TOTAL WEIGHT OF EPICOTYLEDONARY TISSUE IN PLANTS OF TERATOLOGICAL AND NORMAL RACES.—In the foregoing discussion comparisons were limited to the weight of primordial leaves. This was done because of the difficulty of securing leaves subsequently formed in comparable stages of development in the normal and teratological seedlings. It seemed desirable to supplement these studies by the determination of the total weight of tissue produced by the two races. The results for a comparison of the total weight of tissue produced above the cotyledonary node by tetracotyledonous plants and their normal control of line 139, are shown under the heading "epicotyl" in the fundamental tables of data.

The constants show that without exception the green weight and dry weight per plant and percentage dry matter are higher in the normal than in the tetracotyledonous plants. The percentage differences range from -7.6 to -29.4 in the case of green weight, and from -10.1 to -30.5 in the case of dry weight. The differences in percentage of dry matter range from -0.036 to -0.249. The average weight of green tissue per plant is 1.4447 for the abnormal and 1.8088 for the control series, or a difference of -0.3640 gm. The average dry weight per plant is 0.1093 for the abnormal and 0.1384 for the normal seedling, or a difference of -0.0291 gm. The average percentage difference for the green weight is -18.96, while for dry weight the difference is -20.30. The percentage of dry material in the abnormal seedling is 7.6448 as compared with 7.7740 in the control, a difference of -0.1292.

Summary

This paper presents the results of an investigation of green weight, dry weight, and of the ratio of green weight to dry weight in primordial leaf tissue in mutant and parental races of *Phaseolus vulgaris*.   The data show that when grown under as nearly identical conditions as possible the primordial leaves of the mutant (tetra-cotyledonous) show a smaller green weight, a smaller dry weight, and a lower ratio of dry weight to green weight than those of the normal (dicotyledonous) parental race.   Thus the tetracotyle-donous race is distinguished not merely by striking morphological differences, but by physiological differentiation as well.   In this respect the results for the heritable mutant race are in agreement with those for variant individuals within the same strain.

STATION FOR EXPERIMENTAL EVOLUTION
COLD SPRING HARBOR, N.Y.

CORRELATIONS BETWEEN ANATOMICAL CHARACTERS IN THE SEEDLING OF PHASEOLUS VULGARIS

J. ARTHUR HARRIS, EDMUND W. SINNOTT, JOHN Y. PENNYPACKER, AND G. B. DURHAM

(Received for publication January 17, 1921)

INTRODUCTION

In an earlier paper[1] we traced the course of the vascular bundles through-out the dimerous and trimerous seedlings of *Phaseolus vulgaris* and measured the variation occurring at different levels.

The chief results of that paper were (a) the demonstration of the profound differentiation of dimerous and trimerous seedlings in their internal (vascular) as well as in their external characters, (b) the demonstration that the number of bundles at a given level in the seedling is a highly variable rather than a constant character, and (c) that the various organs or regions of the plant body (particularly, in the present case, those which are separated by the vascular anastomoses at the cotyledonary node) differ widely in the magnitude of their variability as to bundle number.

In this paper we propose to consider in quantitative terms the degree of interrelationship between the vascular structures in the different regions of normal and abnormal seedlings. The results of such an investigation will evidently be of considerable morphological interest, since many of the problems of organic form are fundamentally problems of correlation.

Two morphological problems at once present themselves for consideration:

First, is there a high correlation between the vascular topography of two different levels of the same internode, *i.e.*, is the number of vascular bundles constant throughout the length of an internode or is there more or less extensive splitting or anastomosis within the length of such a conventional morphological unit?

Second, is there a definite correlation between the vascular topography below a node and the vascular topography above it, or is the vascular system so fully reorganized at the nodal anastomosis of bundles that, in bundle number, successive internodes are practically independent of one another?

With the present material, these questions may be answered by determining the coefficients of correlation for bundle number between (1) the base and the mid-region of the hypocotyl, and (2) between the various levels of the hypocotyl and the mid-region of the epicotyl. It is these

[1] Harris, J. Arthur, Sinnott, E. W., Pennypacker, John Y., and Durham, G. B. The vascular anatomy of dimerous and trimerous seedlings of *Phaseolus vulgaris*. Amer. Jour. Bot. 8: 63-102. 1921.

problems which we propose first to consider. We shall also compare the normal and abnormal seedlings as to the correlations which they exhibit and shall touch briefly on the problem of the correlation between bundle number in seedlings from the same parent plant.

The frequency distributions of bundle number are in many cases of very narrow range and very skew. There has, therefore, been considerable question as to the formulae to be employed. It has seemed best for various reasons which need not be detailed here, to employ the usual method of product-moment correlation.

PRESENTATION AND ANALYSIS OF DATA

The series of data considered here are in large part the same as those discussed in our earlier paper, but have in some cases been supplemented by the examination of additional sections. These have been included when the dimerous and trimerous seedlings were not true siblings. In lines 75, 93, and 98, the series compared were obtained from the same mothers. In so far as the data are the same as those used earlier, the variation constants for the different characters have already been presented and discussed and require no further comment here. The data from which measures of interrelationship may be computed are given in our fundamental tables A to L. We have, therefore, merely to deduce and discuss the correlation coefficients

Correlation between Bundle Number at Different Levels in the Same Internode

We first turn to the problem of the relationship between the number of bundles—primary double bundles, intercalary bundles, and total bundles— at the base of the hypocotyl and the number in the central region of the hypocotyl. The reader who cares to do so may reconstruct the 24 correlation tables necessary for a consideration of these relationships from our fundamental tables A–L.

TABLE I. *Coefficients of correlation between number of primary double bundles, number of intercalary bundles, and total bundles at base of hypocotyl, and number of bundles in central region of hypocotyl*

Character of Seedlings and Line	N	Correlation for Primary Double Bundles r_{ph}		Correlation for Intercalary Bundles r_{ih}		Correlation for Total Bundles r_{bh}	
Trimerous							
Line 75.......	142	+.378±.049	7.79	+.329±.051	6.51	+.649±.033	19.
Line 93.......	155	+.233±.051	4.55	+.204±.052	3.92	+.469±.042	11.
Line 98.......	183	+.321±.045	7.17	+.253±.047	5.42	+.586±.033	17.
Line 139......	106	+.417±.054	7.71	+.097±.065	1.50	+.531±.047	11.
Line 143......	221	+.556±.031	17.8	+.305±.041	7.40	+.753±.020	38
Dimerous							
Line 75.......	142	+.362±.049	7.35	+.668±.031	21.3	+.797±.021	38
Line 93.......	155	+.641±.032	21.0	+.390±.046	8.50	+.753±.023	32
Line 98.......	183	+.666±.028	24.0	+.555±.035	16.1	+.786±.019	41
Line 139......	305	+.344±.034	10.1	+.898±.008	119.7	+.925±.006	168
Line 143......	420	+.530±.023	22.5	+.634±.020	32.2	+.802±.011	68

DIAGRAM I.   Regression of number of bundles in central region of hypocotyl and in central region of epicotyl on number of primary double bundles at base of hypocotyl in trimerous seedlings. Empirical means represented by solid dots for hypocotyl and by circles for epicotyl.

DIAGRAM 2. Regression of number of bundles in central region of hypocotyl and in central region of epicotyl on number of primary double bundles at base of hypocotyl in dimerous seedlings. Empirical means represented by solid dots for hypocotyl and by circles for epicotyl.

The correlation coefficients between the two classes of bundles which have been recognized at the base of the hypocotyl and the total number of basal bundles (*i.e.*, the sum of the number of primary double bundles and the number of intercalary bundles in the base of the hypocotyl) and the number in the central region of the hypocotyl, appear in table 1.

The correlations are without exception positive in sign and of a material order of magnitude. They have been expressed in terms of regression on diagram 1 for trimerous seedlings and on diagram 2 for dimerous seedlings of the five lines.[2]

Primary Double Bundles and Mid-region of Hypocotyl

The constants showing the relationship between number of primary double bundles and number of bundles in the central region of the hypocotyl, r_{ph}, are shown in the first section of table 1. They are positive and statistically significant in all cases in both dimerous and trimerous seedlings. The average value of the coefficient for the five lines investigated is +.3810 for trimerous seedlings and +.5086 for dimerous seedlings.

Diagram 2 shows that in the case of the normal plants of lines 75, 93, and 143 a straight line represents very well indeed the changes in the mean number of bundles in the hypocotyl with variations in the number of primary double bundles at the base of the hypocotyl. In line 98 the agreement is apparently not so good. This is, however, attributable to the fact that of the 183 plants only two have more than 5 primary bundles. Of these two, one plant is recorded as having 8, which is twice the normal number. In line 139 only plants with two classes of seedlings, those with 4 or 5 primary bundles, are available, and since the regression line must connect the two means it is idle to discuss linearity of regression.

Turning to the trimerous plants represented in diagram 1, we note that because of the small number of plants with other than 5 or 6 primary double bundles the distribution of the empirical means is very irregular indeed. There is some suggestion of non-linearity, but the number of seedlings in the more extreme classes is so small for every line that little stress is to be laid upon them.

In both normal and abnormal plants the slope of the regression line is rather steep, showing a material change in the number of bundles in the central region of the hypocotyl with variations in the number of primary double bundles at the base of the hypocotyl.

Intercalary Bundles and Mid-region of Hypocotyl

The correlation between the number of intercalary bundles and the total number of bundles in the hypocotyl, r_{ih}, are shown in the second

---

[2] The equations on the diagrams show the regression of the number of bundles in the central region of the hypocotyl, H, and in the central region of the epicotyl, E, on the number of primary double bundles, P, at the base of the hypocotyl. The empirical means for the hypocotyl are represented by solid dots, while those of the epicotyl are represented by circles. In both cases the empirical mean number of bundles for the same organ are connected by solid lines when the number of sections averaged was five or more, but by broken lines when the number available was four or less. Fortunately for purposes of graphical representation, the mean number of bundles in both hypocotyl and epicotyl can be drawn on the same diagram. Only the lower lines in each of the five panels of the two diagrams require consideration for the moment.

DIAGRAM 3. Regression of number of bundles in central region of hypocotyl and in central region of epicotyl on number of intercalary bundles at base of hypocotyl in trimerous seedlings. Empirical means represented by solid dots for hypocotyl and by circles for epicotyl.

DIAGRAM 4.   Regression of number of bundles in central region of hypocotyl and in central region of epicotyl on number of intercalary bundles at base of hypocotyl in dimerous plants.   Empirical means represented by solid dots for hypocotyl and by circles for epicotyl.

section of table I. The straight-line equations showing the regression of the number of bundles in the central region of the hypocotyl are recorded and represented graphically on diagram 3 for trimerous seedlings and on diagram 4 for dimerous seedlings. These diagrams, like the two preceding, also give the regression equations and their graphic representation for the epicotyl which will be discussed in a subsequent section.

The correlation coefficients are positive in all cases, and with one exception may be considered statistically significant. They show, however, a considerable irregularity from line to line, presumably because of the varying range and distribution of number of intercalary bundles. The average value of the coefficient is $+.2376$ for trimerous seedlings and $+.6290$ for dimerous seedlings.

Turning to the graphs, we may note that for the dimerous plants the agreements between the empirical and the theoretical means are very good indeed. The slope of the lines for the hypocotyl is very steep.

The graphs for the trimerous plants show far greater irregularities because of the generally small number of the strands but the occasional occurrence of plants with a larger number. Reference to the tables will show that in line 75 there is one seedling with 6 intercalary bundles whereas the remaining 141 seedlings have only 0, 1, or 2 intercalary bundles. In line 93 there is only one seedling with more than 2 intercalary bundles and it has 4. In line 98 all the frequencies with two exceptions fall on 0 or 1 intercalary bundle.

The correlations and equations have been recalculated, leaving these extreme cases out of account. The regression straight lines based on all the material are represented by solid lines. Those in which the extreme class were omitted are represented by broken lines.[3] The removal of these aberrant cases has increased the agreement between the observed and the theoretical means but the fit is still far from satisfactory. The only conclusion which can be drawn from these diagrams is that there is a considerable degree of positive correlation between the number of the intercalary bundles and the number of bundles in the hypocotyl.

Total Basal Bundles and Mid-region of Hypocotyl

The correlations between total bundles (primary double bundles + intercalary bundles) at the base of the hypocotyl and the number of bundles in the central region of the hypocotyl, r_{bh}, are shown in the third section of table I. The straight-line regression equations are given and represented graphically as the lower figures in each panel of diagram 5 for trimerous seedlings and diagram 6 for dimerous seedlings.

As might be expected on *a priori* grounds, these coefficients agree with those for primary double bundles and for intercalary bundles in sign, and

---

[3] For the curtailed series the regression equations are: Line 75, $H = 12.194 + 0.654 I$; Line 93, $H = 12.238 + 0.462 I$; Line 98, $H = 12.030 + 0.473 I$.

DIAGRAM 5.   Regression of number of bundles in central region of hypocotyl and in central region of epicotyl on total number of bundles at base of hypocotyl in trimerous seedlings. Empirical means represented by solid dots for hypocotyl and by circles for epicotyl.

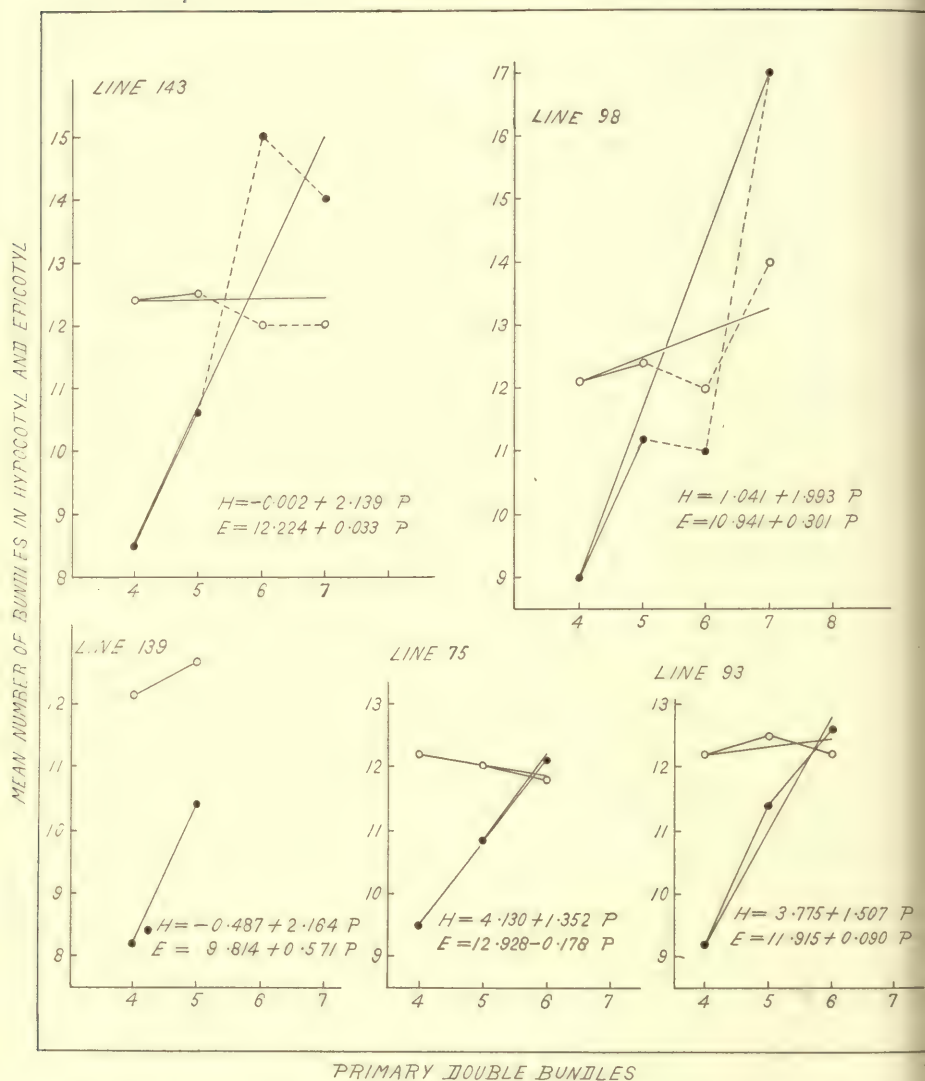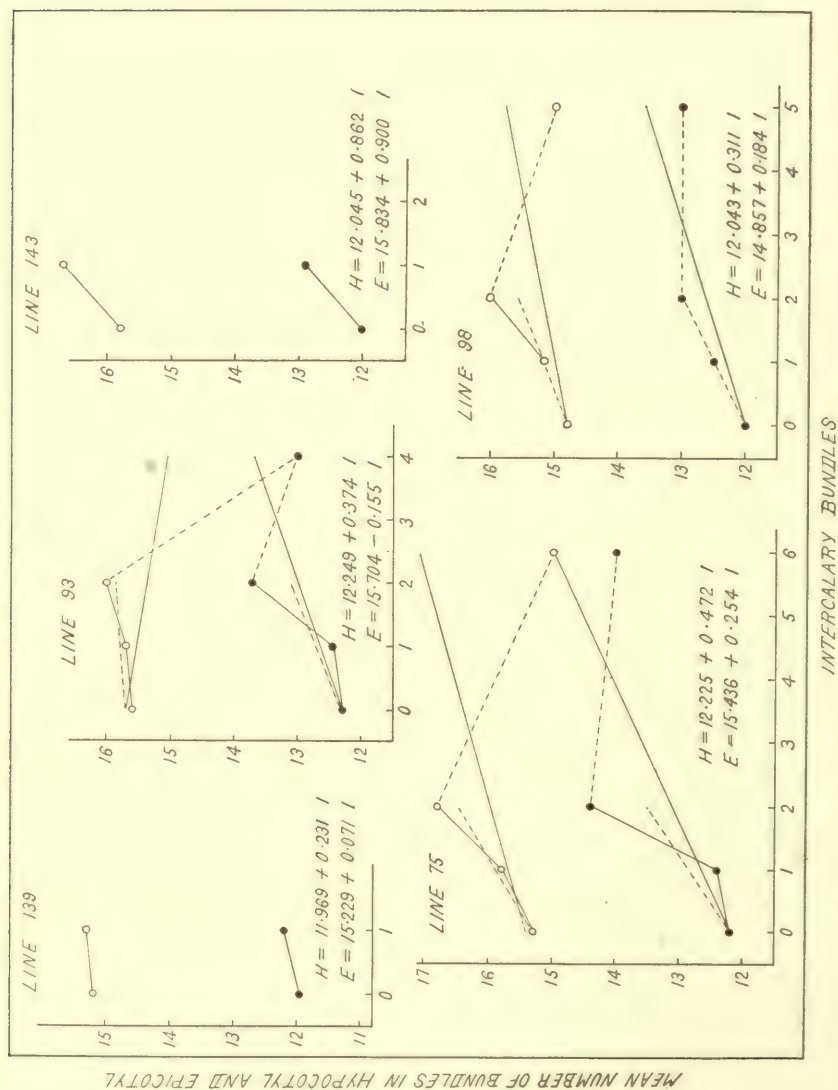DIAGRAM 6. Regression of number of bundles in central region of hypocotyl and in central region of epicotyl on total number of bundles at base of hypocotyl in dimerous seedlings. Empirical means represented by solid dots for hypocotyl and by circles for epicotyl.

.are in general somewhat higher than those for either of these two classes. The average value of the 5 coefficients for trimerous seedlings is $+.5976$ while that for dimerous seedlings is $+.8126$.

Turning to the diagrams, we note that the straight lines and the empirical means are in excellent agreement, considering the small number of seedlings, in the case of the normal plants, but show greater irregularities in the case of the abnormal plants. This is due to a considerable extent to the greater concentration of the frequencies into two classes in the case of the trimerous seedlings.

We may now consider the relative magnitudes of the three correlations which we have been studying. Table 2 shows the differences existing be-

TABLE 2.   *Comparison of correlations between the various types of bundles at the base of hypocotyl and the number of bundles in the central region of hypocotyl*

Character of Seedlings and Line	$r_{bh} - r_{ph}$		$r_{bh} - r_{ih}$		$r_{ph} - r_{ih}$	
Trimerous						
Line 75........	$+.271\pm.059$	4.59	$+.320\pm.061$	5.25	$+.049\pm.071$	0.69
Line 93........	$+.236\pm.066$	3.58	$+.265\pm.067$	3.95	$+.029\pm.073$	0.40
Line 98........	$+.265\pm.056$	4.73	$+.333\pm.057$	5.84	$+.068\pm.065$	1.05
Line 139.......	$+.114\pm.071$	1.61	$+.434\pm.080$	5.43	$+.320\pm.084$	3.81
Line 143.......	$+.197\pm.037$	5.32	$+.448\pm.046$	9.74	$+.251\pm.051$	4.92
Dimerous						
Line 75........	$+.435\pm.053$	8.20	$+.129\pm.037$	3.49	$-.306\pm.058$	5.28
Line 93........	$+.112\pm.040$	2.80	$+.363\pm.051$	7.11	$+.251\pm.056$	4.48
Line 98........	$+.120\pm.033$	3.64	$+.231\pm.040$	5.78	$+.111\pm.045$	2.47
Line 139.......	$+.581\pm.035$	16.6	$+.027\pm.010$	2.70	$-.554\pm.035$	15.8
Line 143.......	$+.272\pm.026$	10.5	$+.168\pm.022$	7.64	$-.104\pm.030$	3.47

tween the various correlations, *i.e.*, the possible differences between the correlation for primary bundles and hypocotyledonary bundles, r_{ph}, for intercalary bundles and hypocotyledonary bundles, r_{ih}, and for total bundles at the base of the hypocotyl and hypocotyledonary bundles, r_{bh}.

For both dimerous and trimerous seedlings, the correlations between the total bundles at the base of the hypocotyl and the number of bundles in the central region of the hypocotyl are higher throughout than those for either of the two separate types of bundles (primary bundles and intercalary bundles) individually considered. In general, the differences are sufficiently large in comparison with their probable errors to be considered statistically significant.

The comparison of the magnitudes of the correlations between numbers of primary double bundles and of vascular elements at higher levels, and between numbers of intercalary bundles and of vascular elements at higher levels, shows that in 7 of the 10 comparisons the closer correlation of hypocotyledonary bundles is with the primary double bundles.

Lines 75, 139, and 143 present exceptions. In the normal plants of these lines the correlation between intercalary bundles and total bundles in the

hypocotyl is apparently significantly higher than that between primary double bundles and total bundles in the hypocotyl.[4]

The fact that the number of bundles in the central region of the hypocotyl is about equally correlated with the number of primary double bundles and with the number of intercalary bundles at the base of the hypocotyl shows that both types of bundles are of about equal significance in determining the number of bundles in the central region of the hypocotyl.

From the foregoing discussion it is clear that there is a rather close relationship between number of bundles at the base and the number in the central region of the hypocotyl. This might, we believe, have been expected on *a priori* morphological grounds. The interesting feature of the results seems to be that the correlations are not larger. The results show that there is a very large amount of irregularity in the division of primary strands or in the formation of intercalary bundles, or in both, as one passes the short distance from the base of the hypocotyl to the central region.

Correlation between Bundle Number in Different Internodes

The data available for a consideration of the problem of the correlation between bundle number in adjacent internodes cover (A) the correlation between the three classes of bundles at the base of the hypocotyl [primary double bundles (p), intercalary bundles (i), and total bundles (b)] and the number of bundles in the central region of the epicotyl; and (B) the correlation between the number of bundles in the central region of the hypocotyl and in the central region of the epicotyl.

(A) The coefficients showing the relationship between the numbers of primary double bundles, of intercalary bundles, and of total bundles at the base of the hypocotyl, and the number of bundles in the central region of the epicotyl, appear in table 3.

The regression equations showing the actual change in number of epicotyledonary bundles associated with variation in the number of primary double bundles are given and are represented with the empirical means of arrays on diagram 1 for trimerous plants and on diagram 2 for dimerous plants.

The graphs for the theoretical lines and the empirical means for the number of bundles in the epicotyl of both normal and abnormal plants show relatively little relationship between the number of bundles at the base of the hypocotyl and the number in the epicotyl. The differences in the slope of the lines for primary basal bundles and the number of bundles in central regions of hypocotyl and epicotyl show in a most striking manner the dif-

---

[4] In line 75 the range of primary double bundles is only 3 while that of intercalary bundles is 6. In line 139 the primary double bundles fall in two classes only, with but 3 of the 305 frequencies on 5 as compared with 302 on 4 bundles. The correlation coefficient in such a case can have but little value. In line 143 practically all of the primary double bundles fall in two classes while the intercalary bundles are limited to three classes.

Irregularity of results must be expected under such conditions.

ferences between correlations for groups of bundles lying on the same side and those lying on different sides of the nodal complex.

(1) The correlation coefficients between primary double bundles and number of bundles in the epicotyl, r_{pe}, as set forth in the first section of table 3, are in part positive and in part negative in sign. For the most part they can not be considered statistically significant. The average value of those for trimerous seedlings is $-.0226$ while that for dimerous seedlings is $+.0768$.

(2) For the correlation between the number of intercalary bundles and the number of bundles in the epicotyl, r_{ie}, shown in the second section of table 3, the coefficients are not in general certainly significant in com-

TABLE 3. *Coefficients of correlation between number of primary double bundles, number of intercalary bundles and total number of bundles at base of hypocotyl, and number of bundles in central region of epicotyl*

Character of Seedlings and Line	N	Correlation for Primary Double Bundles r_{pe}		Correlation for Intercalary Bundles r_{ie}		Correlation for Total Bundles r_{be}	
Trimerous							
Line 75........	142	$+.053\pm.056$	0.93	$+.126\pm.056$	2.27	$+.182\pm.055$	3.33
Line 93........	155	$-.087\pm.054$	1.61	$-.055\pm.054$	1.01	$-.148\pm.053$	2.79
Line 98........	183	$+.008\pm.050$	0.70	$+.070\pm.050$	1.42	$+.099\pm.049$	2.01
Line 139.......	106	$-.105\pm.064$	1.63	$+.016\pm.065$	0.25	$-.095\pm.065$	1.47
Line 143.......	221	$+.018\pm.045$	0.40	$+.233\pm.043$	5.44	$+.190\pm.044$	4.34
Dimerous							
Line 75........	142	$-.115\pm.050$	2.07	$-.043\pm.057$	0.75	$-.054\pm.056$	0.96
Line 93........	155	$+.084\pm.054$	1.55	$+.132\pm.053$	2.48	$+.167\pm.053$	3.16
Line 98........	183	$+.239\pm.047$	5.08	$+.109\pm.049$	2.21	$+.205\pm.048$	4.29
Line 139.......	305	$+.164\pm.038$	4.37	$+.145\pm.038$	3.84	$+.175\pm.037$	4.68
Line 143.......	420	$+.012\pm.033$	0.38	$+.134\pm.032$	4.15	$+.121\pm.032$	3.73

parison with their probable errors. Two of the ten are indeed negative in sign. The coefficients for line 143 in both trimerous and dimerous seedlings and possibly that for line 139 in the dimerous seedlings may be significant. The fact that eight of the ten coefficients are positive suggests that there is a slight relationship between the number of intercalary bundles at the base of the hypocotyl and the number of vascular elements in the central region of the epicotyl. The general average is $+.0780$ for the trimerous and $+.0954$ for the dimerous.

This suggestion is only slightly strengthened by inspection of the two sets of diagrams on which the regression equations are presented and drawn with the empirical means. Diagram 3 pictures the results for trimerous seedlings while the comparable representations for dimerous seedlings are shown on diagram 4. These show that while the slope showing the change in the number of bundles in the hypocotyl associated with variations in the number of intercalary bundles at the base of the epicotyl is very steep, it is practically nothing for the epicotyl, thus indicating a very close relationship in the former case but the practical absence of interdependence in the latter.

As explained above (p. 346), the slopes for the trimerous seedlings are very greatly influenced by certain aberrant individuals. When these are removed we obtain the equations represented by the broken lines in the figures.[5] The results for the relationship between the number of intercalary bundles and the number of bundles in the epicotyl indicate a positive correlation in all 3 cases when the one extreme plant is removed.

(3) The coefficients of correlation between total bundles (double bundles plus intercalary bundles) at the base of the hypocotyl and the number of bundles in the central region of the epicotyl, r_{be}, are shown in the third section of table 3, and are represented graphically in terms of regression in the upper figures of each panel of diagram 5 for trimerous seedlings and of diagram 6 for normal seedlings. The very gentle slope and the differences in direction of the lines for the epicotyl of the trimerous plants, together with the irregularity of the empirical means, serve to emphasize the slightness of the relationship between total bundles at the base of the hypocotyl and the number of bundles in the central region of the epicotyl. In the normal plantlets the means are less irregularly distributed about the theoretical lines, but the slope of the lines is very slight, and in one case the regression slope has the negative sign.

Turning to the correlation constants for more direct numerical comparison, we note that three of the ten constants are negative. The general average is $+.0456$ for the trimerous and $+.1228$ for the dimerous seedlings.

Looking back over diagrams 1–6, one cannot but be impressed by the difference in the slope of the lines showing the changes in number of bundles in the hypocotyl and in the epicotyl respectively associated with variations in the number of bundles at the base of the hypocotyl. The lines for the hypocotyl, without exception, indicate an increase in the number of bundles in the central region of the hypocotyl with an increase in the number of bundles at the base of the hypocotyl. The lines for the epicotyl occasionally show a decrease. Furthermore, the slopes of the lines for the hypocotyl are in general conspicuously steeper—thus indicating closer dependence upon the number of basal bundles—than those for the epicotyl.

Turning to table 4 for a numerical comparison of the correlations between the systems of bundles on the same side and on different sides of the cotyledonary node, we note that without exception the coefficients of correlation measuring the interrelationship between the number of vascular elements at the base of the hypocotyl and in the central region of the epicotyl are markedly lower than those measuring the correlation between the number of vascular elements in the base of the hypocotyl and in the central region of the hypocotyl.

(B) We now have to consider the problem of the correlation between the numbers of bundles in the central regions of the hypocotyl and of the

---

[5] When the extreme cases are omitted the equations are: Line 75, $E = 15.378 + 0.591\ I$; Line 93, $E = 15.670 + 0.096\ I$; Line 98, $E = 14.840 + 0.394\ I.$

TABLE 4. *Differences between correlations for three classes of bundles at base of hypocotyl and the number of bundles in the central regions of hypocotyl and epicotyl, respectively*

Character of Seedlings and Line	$r_{pe} - r_{ph}$		$r_{ie} - r_{ih}$		$r_{be} - r_{bh}$	
Trimerous						
Line 75........	$-.325 \pm .074$	4.39	$-.203 \pm .075$	2.71	$-.467 \pm .064$	7.29
Line 93........	$-.320 \pm .074$	4.32	$-.259 \pm .075$	3.45	$-.617 \pm .068$	9.07
Line 98........	$-.313 \pm .067$	4.67	$-.183 \pm .069$	2.65	$-.487 \pm .059$	8.25
Line 139.......	$-.522 \pm .084$	6.21	$-.081 \pm .092$	0.88	$-.626 \pm .080$	7.83
Line 143.......	$-.538 \pm .055$	9.78	$-.072 \pm .059$	1.22	$-.563 \pm .048$	11.7
Dimerous						
Line 75........	$-.477 \pm .070$	6.81	$-.711 \pm .065$	10.9	$-.841 \pm .059$	14.3
Line 93........	$-.557 \pm .062$	8.98	$-.258 \pm .070$	3.69	$-.586 \pm .057$	10.3
Line 98........	$-.427 \pm .055$	7.76	$-.446 \pm .060$	7.43	$-.581 \pm .052$	11.2
Line 139.......	$-.180 \pm .051$	3.53	$-.753 \pm .039$	19.3	$-.750 \pm .037$	20.3
Line 143.......	$-.518 \pm .040$	12.9	$-.500 \pm .037$	13.5	$-.681 \pm .033$	20.6

epicotyl of the plant. The correlation surfaces are given in tables A–L. The results are set forth in table 5.

TABLE 5. *Coefficient of correlation between number of bundles in central region of hypocotyl and central region of epicotyl*

Line	Trimerous			Dimerous		
	N	r	r/E_r	N	r	r/E_r
75.....	416	$+.012 \pm .033$	0.36	416	$-.017 \pm .033$	0.52
93.....	557	$+.075 \pm .028$	2.68	557	$+.162 \pm .028$	5.79
98.....	345	$+.090 \pm .036$	2.50	345	$+.225 \pm .035$	6.43
139.....	106	$-.061 \pm .065$	0.94	305	$-.187 \pm .037$	5.05
143.....	143	$+.256 \pm .042$	6.10	420	$+.107 \pm .033$	3.24

The correlations are positive with the exception of that for dimerous plants of line 75 and of that for both dimerous and trimerous plants of line 139, which are negative in sign. Only one of the negative coefficients may be considered statistically significant in comparison with its probable error. Several of the positive coefficients are large enough in comparison with their probable errors to be considered possibly significant. The average correlation for the trimerous plants is $+.074$ while that for the dimerous plants is $+.058$. The correlations for the trimerous and dimerous plants can not be considered to differ significantly.

The generally positive sign of the constants suggests that seedlings which have a larger number of bundles in the hypocotyl have on the average a larger number of bundles in the epicotyl. This is the condition actually found in the series studied, but the difficulties in the interpretation of the probable error in cases in which the correlation coefficient is so small should make one cautious in generalizing the results obtained.

How slight the relationship between the numbers of bundles in the two organs is, may be shown by the regression lines giving the change in the mean number of bundles in the epicotyl associated with variations in the

number of bundles in the hypocotyl and in the mean number of bundles in the hypocotyl associated with variations in the epicotyl. The straight line equations are as follows:

	Dimerous	Trimerous
Line 75,	$H = 10.325 - .068E$	$H = 12.055 + .009E$
	$E = 12.347 - .008H$	$E = 15.267 + .016H$
Line 93,	$H = 5.736 + .401E$	$H = 11.501 + .050E$
	$E = 11.494 + .065H$	$E = 14.273 + .112H$
Line 98,	$H = 1.374 + .648E$	$H = 11.408 + .042E$
	$E = 11.388 + .078H$	$E = 12.538 + .195H$
Line 139,	$H = 4.105 + .338E$	$H = 12.492 - .033E$
	$E = 11.254 + .103H$	$E = 16.591 - .113H$
Line 143,	$H = 6.677 + .161E$	$H = 9.279 + .187E$
	$E = 11.737 + .072H$	$E = 11.810 + .349H$

All of these lines have been drawn, but it seems unnecessary to publish more than three sets.

The comparison between the empirical and the theoretical mean number of bundles in the epicotyls of seedlings classified according to the number of bundles in the hypocotyl is made for three lines on diagram 7. Conversely, the comparison of the actual mean number of bundles in the hypocotyl for plants with various numbers of bundles in the epicotyl is made on diagram 8.

The slight slope of the lines and the irregularity of the empirical means show in a very convincing manner the laxness of the relationship between the numbers of bundles in the central regions of hypocotyl and epicotyl.

These results are of decided morphological significance. The profound difference between the correlations for the hypocotyl and for the epicotyl emphasizes the completeness of the loss of individuality of the bundles at the cotyledonary node. Whereas the number of bundles in the central region of the hypocotyl is quite closely correlated with the number at the base of the hypocotyl, there cannot be asserted to be any significant correlation in bundle number between either the base or the central region of the hypocotyl and the central region of the epicotyl, when we deal with seedlings of the same gross morphological structure. In other words, the reorganization of the vascular system at the node is so complete that the portion of the system which is above the node shows practically no relation to the portion which is below the node.

Comparison of Correlation in Trimerous and Dimerous Seedlings.

In examining the results of the preceding tables the reader may have noted that the coefficients for the dimerous are preponderantly higher than those for the trimerous plants. This result is clearly brought out in table 6

DIAGRAM 7.   Regression of number of bundles in central region of epicotyl on number of bundles in central region of hypocotyl.   Empirical means represented by solid dots for dimerous seedlings and by circles for trimerous seedlings.

in which the differences between the coefficients for the two classes of plants are shown.

The differences in this table are generally negative, thus indicating that the correlations are lower in the trimerous than in the dimerous seedlings. The exceptions are of some interest.

DIAGRAM 8. Regression of number of bundles in central region of hypocotyl on number of bundles in central region of epicotyl. Empirical means represented by solid dots for dimerous seedlings and by circles for trimerous seedlings.

There are only 4 exceptions among the 15 correlations between the numbers of vascular elements in the basal region of the hypocotyl and in the central region of the hypocotyl, as shown in the upper section of the table. These are without exception insignificant in comparison with their probable errors. There are 9 exceptions among the 20 correlations be-

TABLE 6. *Comparison of correlations for trimerous and dimerous seedlings. Differences only (trimerous less dimerous) are given. See tables 1 and 3 for constants*

	Correlation Coefficient Compared					
	r_{ph}		r_{ih}		r_{bh}	
Line 75	+.016±.069	0.23	−.339±.060	5.65	−.148±.039	3.79
Line 93	+.408±.060	6.80	−.186±.069	2.69	−.284±.048	5.92
Line 98	−.345±.053	6.50	−.302±.058	5.21	−.200±.039	5.13
Line 139	+.073±.064	1.14	−.801±.066	12.1	−.394±.047	8.38
Line 143	+.026±.039	0.67	−.329±.046	7.15	−.049±.022	2.23
	r_{pe}		r_{ie}		r_{be}	
Line 75	+.168±.075	2.24	+.169±.080	2.11	+.236±.079	2.99
Line 93	−.171±.076	2.25	−.187±.075	2.49	−.315±.075	4.20
Line 98	−.231±.069	3.35	−.039±.070	5.57	−.106±.069	1.53
Line 139	−.269±.074	3.64	−.129±.075	1.72	−.270±.075	3.60
Line 143	+.006±.056	0.11	+.099±.054	1.83	+.069±.055	1.25
	r_{he}					
Line 75	+.029±.047	0.62	—	—	—	—
Line 93	−.087±.040	2.18	—	—	—	—
Line 98	−.135±.050	2.70	—	—	—	—
Line 139	+.126±.075	1.68	—	—	—	—
Line 143	+.149±.053	2.81	—	—	—	—

tween the numbers of vascular elements on different sides of the cotyledonary node as shown in the central and lower section.   The exceptions occur, in short, among the relationships which in both types of seedlings are practically zero in intensity.

We have no explanation to offer of this greater intensity of correlation in the sub-cotyledonary region of the normal seedling.   The result is stated as one of the matters of fact demonstrated by the investigation.

Correlation between Bundle Number in Siblings

The question will naturally arise as to whether the variability in number of bundles in both hypocotyl and epicotyl and the correlation between bundle number in these two internodes may be due to a differentiation of the parent plants from which the seeds were obtained, either in their genetic composition or because of environmental influences.   This problem presents many difficulties.   Some light may be thrown upon it in the following manner.

An abnormal and a normal seedling were taken from the same parent plant.   Thus it is possible to determine in our series the correlation between the number of bundles in the hypocotyl of an abnormal plant and in the hypocotyl of a normal plant derived from the same parent.   If a differentiation of the parent plants due to either genetic or physiological factors is the underlying proximate cause of the variability and correlation in bundle number in seedlings which we have studied, there should be a correlation between the number of bundles in the seedlings derived from the same plant.

The correlations between the numbers of bundles in the hypocotyls

and epicotyls of the normal and abnormal seedling, *i.e.*, of dimerous and trimerous seedlings, from the same parent plants are given in table 7.[6]

TABLE 7.   *Correlations between bundle number in offspring of same parent plant*

Character of Plant and Organs Compared		Line and Correlation		
Trimerous	Dimerous	Line 75	Line 93	Line 98
Hypocotyl	Hypocotyl			
	C. S. H.....	+.0540±.0406	+.1703±.0327	−.0512±.0529
	Storrs......	+.2151±.0540	+.0553±.0540	+.0853±.0495
Epicotyl......	Epicotyl			
	C. S. H.....	−.0037±.0407	−.0027±.0336	+.1222±.0522
	Storrs......	+.0685±.0563	+.0432±.0541	+.0401±.0498

The coefficients are low throughout.   Nine of the 12 are positive while 3 are negative in sign.   Only 2 of the 12 can be reasonably regarded as significant.   Both of these are positive.   There is, therefore, a suggestion of a positive correlation between the anatomical characters of seedlings from the same parent.   The values are too low, however, to justify the conclusion that there is a measurable differentiation in the genetic or physiological characteristics of the parent plants affecting bundle number in the offspring seedling.

The absence of correlation here connotes an absence of (sororal or fraternal) inheritance in bundle number.

SUMMARY

In an earlier paper we have shown that the number of vascular elements at different levels in the seedling of *Phaseolus vulgaris* is subject to considerable variation and that the amount of variation may itself differ from level to level.   This is true both in normal seedlings with two cotyledons and two primordial leaves and in variant seedlings with three cotyledons and a whorl of three primordial leaves.   These two types of seedlings are profoundly differentiated in vascular anatomy as well as in superficial structure.

The purpose of the present paper is to consider the correlations between the number of bundles in the various regions of the seedling.   The characters considered are (1) number of primary double bundles, of intercalary bundles, and of total bundles at the *base* of the hypocotyl, (2) number of bundles in the *central region* of the hypocotyl, and (3) number of bundles in the central region of the epicotyl.

1. There is a substantial correlation between each of the three classes of bundles at the base of the hypocotyl and the number of bundles in the central region of the hypocotyl.   In the normal seedlings the coefficients

---

[6] It has not seemed worth while to publish the tables upon which these very slight correlations are based.   For purposes of comparison the series sectioned at Cold Spring Harbor and at Storrs are both given.

average +.509 for primary double bundles and hypocotyledonary bundles, -.629 for intercalary bundles and hypocotyledonary bundles, and +.813 or total bundles and hypocotyledonary bundles. In the trimerous plants these correlations average +.381, +.238, and +.598, respectively.

The correlations for normal plants are generally higher than those for abnormal plants.

2. The correlation between each of the three classes of bundles at the base of the hypocotyl and the number of bundles in the central region of the epicotyl is low. The coefficients are sometimes positive and sometimes negative in sign. On the basis of the data available it is impossible to assert that there is any correlation at all between the numbers of bundles in these two regions.

3. The correlation between the numbers of bundles in the central region of the hypocotyl and in the central region of the epicotyl is likewise very low. The coefficients are generally not significant in comparison with their probable errors. If there be any correlation at all between the numbers of bundles in these two regions it is very slight indeed.

These results for correlation fully substantiate the conclusions drawn in an earlier paper that there is a *complete reorganization of the vascular system at the cotyledonary node.*

4. The correlation between the number of bundles (either hypocotyledonary or epicotyledonary) in siblings is, if it exists at all, very low. The differentiation of the parent plants through either genetic or environmental factors cannot, therefore, be considered to be the source of the variation and correlation in bundle number demonstrated in this and in our preceding paper.

Conclusions

These results, and others for which the reader must turn back to the body of the paper, justify the emphasis at this point of the following general conclusions:

a. The vascular structures of the seedling are not constant but are decidedly variable within the species. They show different degrees of variability within the individual organism.

b. Seedlings differing in external form are profoundly differentiated in their internal anatomy. This differentiation is evident both in mean number of bundles and in the degree of variability in bundle number. In short, the external form and the internal structure of the seedling are highly but not perfectly correlated.

c. The different anatomical characters of the seedling are interrelated with varying degrees of intensity. Between some there is a very strong correlation, but between others practically none at all.

The quantitative measurement and interpretation of such relationships, by means of the biometric methods hitherto little applied in the field of vascular morphology, will make possible material advance in the investigation of the fundamental problems of morphogenesis.

TABLE A. *Data for correlation between bundle number at the base of the hypocotyl and in the central regions of hypocotyl and epicotyl in trimerous seedlings*

Base*	Line	Hypocotyl													Epicotyl											Totals
		8	9	10	11	12	13	14	15	16	17	18	19	20	12	13	14	15	16	17	18	19	20	21	22	
(4)+0...	143	2																1		1						2
(4)+1...	139			1												1										1
	143		1		1															3						3
(4)+4...	93						1									1										1
(4)+5...	98						1											1								1
(4)+6...	75							1										1								1
(5)+0...	75					1													1							1
	93			1	1	3												1	1	2		1				5
	98			3		1										1	3									4
	139			1	2	1										1		1		1		1				4
	143			11	2	1				1					2		3	3	3	1		3				15
(5)+1...	75			2	2	3	1										2	3	1	2						8
	93				6	3		1							1		1	1	3	1	2	1				10
	98			1	3	2										1	2	2	1							6
	139			1	1	1	1										1	1		2						4
	143				11	13	2	4			1					2	2	9	6	5	3	1	2	1		31
(5)+2...	75						1			1							1			1						2
	93						1	2									1			1	1					3
	98						1												1							1
(6)+0...	75				6	90	8	3								5	19	43	22	11		6		1		107
	93				2	102	10	4	1					1		6	14	47	25	15	8	3	1		1	120
	98				4	145	8	3							6	13	36	67	26	8	3		1			160
	139			1	5	81	3	1	1							4	21	35	21	8	2	1				92
	143					121	7	4	1	1	4				3	7	14	34	31	22	19	3	1			134
(6)+1...	75					1	9	2										6	3	1		1	1			12
	93						8	3							1			6	2	1	1					11
	98						7	3									3	3	2		1	1				10
	139					1	2	2								2		1	2							5
	143					1	11	12	1								5	5	8	3	2	2				25
(6)+2...	75					1			1													2				2
	93							1										1								1
(7)+0...	75							7										4	2	1						7
	93							3	1								2		1	1						4
	98						1									1										1
	143							5									2	2	1							5
(7)+1...	143								3	1							1	1		1				1		4
(7)+2...	75							1								1										1
(8)+0...	75									1									1							1
	143							1													1					1
(8)+1...	143											1										1				1

* Numbers in parentheses are of *primary double bundles*; those following are of *intercalary bundles*.

Table B. *Data for correlation between bundle number at the base of the hypocotyl and in the central regions of hypocotyl and epicotyl in dimerous seedlings*

Base	Line	Hypocotyl										Epicotyl							Tot.
		8	9	10	11	12	13	14	15	17	18	10	11	12	13	14	15	16	
(4)+0	75	40	17	7	4			I				I	I	59	4		2	2	69
	93	12	13	6	3									29	4	I			34
	98	57	31	9										88	8			I	97
	139	269	I											250	18	2			270
	143	262	21	6	2									229	43	13	4	2	291
(4)+1	75		14	13	3								2	22	3	2		I	30
	93		14	17	3	I	I	I						31	3	3			37
	98	2	26	13	I	I								37	5	I			43
	139		22	3	I									21	4	I			26
	143	I	61	27	10		2		I					72	17	11	I	I	102
(4)+2	75			2	4	3	I							9	I				10
	93			9	2	2								11	2				13
	98		3	13	6	I								19	3		I		23
	139			2	2	I	I							5	I				6
	143			3	I	I								3		I		I	5
(4)+3	75				2	I				I				4					4
	93				3	2								3		I	I		5
	98				2									2					2
(4)+4	75							I		I				2					2
	93						I							I					1
(4)+5	75							I		I				2					2
(5)+0	75		I	8	2	I	I							12	I				13
	93		I	13	4	3	I							18	I	2	I		22
	98			2	3	I								4	I	I			6
	139			I										I					1
	143		I	11		I								8	5				13
(5)+1	75				2	2								4					4
	93			I	6	7	3	I						12	3	I	2		18
	98			I	5	I	I							7	I				8
	139			I	I									I		I			2
	143				4	2	I							4	I	2			7
(5)+2	75					I								I					1
	93					5	3	I						6	I	2			9
	98					I								I					1
(5)+3	75						I							I					1
	93						I							I					1
	98						I									I			1
(6)+0	75				I	3		I					I	4					5
	93				I	8	I							9	I				10
	98				I									I					1
(6)+1	75					I								I					1
	93						I	2						2	I				3
	143							I						I					1
(6)+2	93							2						I		I			2
(7)+0	143						I							I					1
(8+1)	98										I					I			

TABLE C. *Correlation between numbers of bundles in hypocotyl and epicotyl of trimerous plants of line 75*

Hypocotyl	Epicotyl										Totals
	12	13	14	15	16	17	18	19	20	21	
8......	......	......	1	......	......	......	......	......	......	......	1
9......	......	......	1	1	......	......	......	......	......	1	3
10......	......	......	......	5	......	......	......	......	......	......	5
11......	1	......	6	12	12	3	2	......	......	......	36
12......	2	15	48	120	62	21	19	3	2	......	292
13......	......	......	5	13	10	7	4	1	......	......	40
14......	......	......	1	10	7	8	2	......	1	......	29
15......	......	......	1	2	......	1	......	......	1	......	5
16......	......	......	......	......	1	......	......	......	......	......	1
17......	......	1	......	1	1	1	......	......	......	......	4
Totals	3	16	63	164	93	41	27	4	4	1	416

TABLE D. *Correlation between numbers of bundles in hypocotyl and epicotyl of dimerous plants of line 75*

Hypocotyl	Epicotyl							Totals
	10	11	12	13	14	15	16	
8..........	......	......	116	17	6	4	......	143
9..........	1	2	78	13	5	2	2	103
10..........	......	......	74	10	......	1	1	86
11..........	......	1	29	4	2	2	......	38
12..........	......	1	22	1	1	1	......	26
13..........	......	......	6	1	......	......	......	7
14..........	......	......	8	......	1	......	......	9
15..........								
16..........								
17..........	......	......	2	......	1	......	......	3
18..........	......	......	1	......	......	......	......	1
Totals.....	1	4	336	46	16	10	3	416

TABLE E. *Correlation between numbers of bundles in hypocotyl and epicotyl of trimerous plants of line 93*

Hypocotyl	Epicotyl											Totals
	12	13	14	15	16	17	18	19	20	21	22	
10............	......	......	......	3	1	1	3	......	......	......	......	8
11............	......	......	3	14	7	4	3	1	......	......	......	32
12............	3	14	33	170	92	38	26	5	1	......	......	382
13............	1	4	5	30	15	6	16	3	1	......	1	82
14............	1	......	5	17	9	2	3	1	......	......	......	38
15............	......	......	1	2	5	2	......	......	2	......	......	12
16............	......	......	......	......	......	1	......	......	......	......	......	1
17............	......	......	......	......	......	......	......	......	......	......	......	
18............	......	......	......	......	......	1	......	......	......	......	......	1
19............	......	......	......	......	......	......	......	......	......	......	......	
20............	......	......	......	......	......	1	......	......	......	......	......	1
Totals........	5	18	47	236	129	56	51	10	4	......	1	557

TABLE F. *Correlation between numbers of bundles in hypocotyl and epicotyl of dimerous plants of line 93*

Hypocotyl	Epicotyl							Totals
	10	11	12	13	14	15	16	
8...............	......	1	31	2	......	......	......	34
9...............	1	1	85	3	3	......	......	93
0...............	......	3	147	14	3	2	......	169
1...............	......	1	88	14	2	......	......	105
2...............	......	......	78	5	9	3	1	96
3...............	......	......	31	4	......	4	......	39
4...............	......	......	16	......	1	1	......	18
5...............	......	......	1	......	......	......	......	1
8...............	......	......	2	......	......	......	......	2
Totals........	1	6	479	42	18	10	1	557

TABLE G. *Correlation between numbers of bundles in hypocotyl and epicotyl of trimerous plants of line 98*

Hypocotyl	Epicotyl										Totals
	12	13	14	15	16	17	18	19	20	21	
9.......	......	......	1	......	......	......	......	......	......	......	1
10.......	......	1	1	4	......	......	......	......	......	......	6
11.......	......	1	2	4	3	1	1	......	......	......	12
12.......	8	20	58	157	41	8	4	......	1	......	297
13.......	......	2	6	8	1	......	2	1	......	1	21
14.......	......	......	1	3	4	......	......	......	......	......	8
Totals .	8	24	69	176	49	9	7	1	1	1	345

TABLE H. *Correlation between numbers of bundles in hypocotyl and epicotyl of dimerous plants of line 98*

Hypocotyl	Epicotyl					Totals
	12	13	14	15	16	
8...............	107	6	......	......	......	113
9...............	103	6	......	......	1	110
10...............	71	4	1	1	......	77
11...............	26	5	1	......	......	32
12...............	7	2	......	......	......	9
13...............	2	......	1	......	......	3
17...............	......	......	1	......	......	1
Totals........	316	23	4	1	1	345

TABLE I. *Correlation between numbers of bundles in hypocotyl and epicotyl of trimerous plants of line 139*

Hypocotyl	Epicotyl							Totals
	13	14	15	16	17	18	19	
10.............	2	......	......	......	1	1	......	4
11.............	1	1	2	1	2	......	1	8
12.............	4	19	32	20	5	3	1	84
13.............	......	1	2	2	1	......	......	6
14.............	1	......	2	......	......	......	......	3
15.............	......	......	......	1	......	......	......	1
Totals........	8	21	38	24	9	4	2	106

TABLE J. *Correlation between numbers of bundles in hypocotyl and epicotyl of dimerous plants of line 139*

Hypocotyl	Epicotyl			Totals
	12	13	14	
8..................	249	18	2	269
9..................	20	3	......	23
10.................	5	1	1	7
11.................	2	1	1	4
12.................	1	......	......	1
13.................	1	......	......	1
Totals............	278	23	4	305

TABLE K. *Correlation between numbers of bundles in hypocotyl and epicotyl of trimerous plants of line 143*

Hypocotyl	Epicotyl										Totals
	12	13	14	15	16	17	18	19	20	21	
8.........	......	......	......	......	1	......	1	......	......	......	2
9.........	......	......	......	......	......	......	1	......	......	......	1
10.........	2	......	2	2	2	1	......	2	......	......	11
11.........	......	......	2	2	5	2	1	1	1	......	14
12.........	3	9	14	35	31	22	19	3	......	......	136
13.........	......	......	1	6	5	5	2	1	1	......	21
14.........	......	......	......	7	4	6	4	1	2	1	25
15.........	......	......	......	2	......	......	3	......	......	1	6
16.........	......	......	......	......	1	......	......	1	1	......	3
17.........	......	......	......	......	......	1	......	......	......	......	1
18.........	......	......	......	......	......	......	......	......	1	......	1
Totals....	5	9	19	54	49	37	31	9	6	2	221

TABLE L.   *Correlation between numbers of bundles in hypocotyl and epicotyl of dimerous plants of line 143*

Hypocotyl	Epicotyl					Totals
	12	13	14	15	16	
8...............	205	42	10	4	2	263
9...............	65	11	7	....	....	83
10...............	32	8	5	....	2	47
11...............	9	3	4	1	....	17
12...............	3	....	1	....	....	4
13...............	1	2	....	....	....	3
14...............	1	....	....	....	....	1
15...............	2	....	....	....	....	2
Totals........	318	66	27	5	4	420

HE VASCULAR ANATOMY OF HEMITRIMEROUS SEEDLINGS OF PHASEOLUS VULGARIS

J. Arthur Harris, Edmund W. Sinnott, John Y. Pennypacker, and G. B. Durham

INTRODUCTORY

In an earlier paper[1] we discussed the gross vascular anatomy of dimerous nd trimerous seedlings of the garden bean. By *dimerous* seedlings we nderstand those of the normal type, characterized by two cotyledons and wo primordial leaves, both sensibly opposite in insertion. By *trimerous* e mean those which have a whorl of three cotyledons and three primordial aves. The cotyledons may be, and frequently are, more or less irregular a insertion. The primordial leaves are, in the seedlings considered, inserted a a regular whorl.

In addition to these two types of seedlings, those which are in a sense atermediate in superficial structure between the two types hitherto studied aay occur. These are seedlings with a whorl of three cotyledons but with normal pair of primordial leaves instead of three as in the case of trimerous eedlings. These we have called *hemitrimerous*. They are extremely rare a occurrence, but during the four years during which these studies have een under way a number sufficiently large to justify a brief discussion of heir gross vascular anatomy has been secured.

Our purpose in this paper is to compare the anatomy of these hemi- rimerous seedlings with the trimerous seedlings (in common with which hey have three cotyledons) on the one hand and with dimerous seedlings in common with which they have two primordial leaves) on the other.

For convenience of reference the three types will in some cases be lesignated by the number of cotyledons and primordial leaves: 2–2 = dim- rous, 3–3 = trimerous, and 3–2 = hemitrimerous.

MATERIALS

The hemitrimerous plants and the trimerous and dimerous seedlings vith which they are compared were largely secured in the series of germina-

[1] Harris, J. Arthur, Sinnott, E. W., Pennypacker, J. Y., and Durham, G. B. The 'ascular anatomy of dimerous and trimerous seedlings of *Phaseolus vulgaris*. Amer. our. Bot. 8: 63–102. 1921.

tions which furnished the materials for our earlier discussion of dimerous and trimerous seedlings. The dimerous and hemitrimerous seedlings were derived from the same parent plants in lines 75, 93, and 98. In lines 29, 139, and 143 the germinations were made from mass seed instead of from the seed of individual parent plants. All of the seed, however, was grown in the same experimental field in 1917.

Since it has been shown in an earlier paper[2] that there is practically no correlation between the anatomical characters of the trimerous and dimerous seedlings from the same parent plant, we are fully justified in using random samples of hemitrimerous, trimerous, and dimerous seedlings for a comparison of their vascular characters.

A detailed account of the vascular topography of the dimerous and trimerous seedling is presented in a previous paper by the writers, but may be summarized very briefly here. Each primary polar bundle of the root bifurcates in the base of the hypocotyl to form a "primary double bundle," which gives rise to two distinct and well separated strands in the central region of the hypocotyl. In addition to these, there are usually present in the hypocotyl a number of "intercalary" bundles, arising either *de novo* or by splitting of some of the primary strands. At the cotyledonary node a rather complex vascular anastomosis takes place, from which the cotyledonary strands depart and out of which the vascular system of the epicotyl is organized.

PRESENTATION AND ANALYSIS OF STATISTICAL DATA

Base of Hypocotyl

The frequency distribution of the various types of vascular organization at the base of the hypocotyl is shown for all the available data in table 1. In this table the number of primary double bundles appears in parentheses, while the number of intercalary bundles follows the + sign.

Because of the relatively small numbers of hemitrimerous seedlings which can be obtained and because of the irregularity of the frequency distributions for bundle number, it has not seemed desirable in this paper to consider the frequency distributions of the numbers of bundles of the several types. Neither has it seemed desirable, on the basis of the relatively small series of hemitrimerous seedlings which can be obtained, to consider the relative variabilities of bundle number in the different regions of the three types of seedlings as we did in our discussion of variation in the dimerous and trimerous types. We have, therefore, limited ourselves to a comparison of mean bundle number, leaving the question of variability until larger series of countings can be obtained.

[2] Harris, J. Arthur, Sinnott, E. W., Pennypacker, J. Y., and Durham, G. B. Correlations between anatomical characters in the seedling of *Phaseolus vulgaris*. Amer. Jour. Bot. **8**: 339–365. 1921.

TABLE I.

Base of hypocotyl	Line 29			Line 75			Line 98			Line 139			Line 143		
	3–3	3–2	2–2	3–3	3–2	2–2	3–3	3–2	2–2	3–3	3–2	2–2	3–3	3–2	2–2
.) + 0..	—	I	83	—	—	101	—	—	117	—	2	270	2	3	291
;) + 1..	—	4	11	—	2	39	—	I	55	I	5	26	3	6	102
) + 2..	I	2	I	—	I	14	—	I	32	—	—	6	—	—	5
) + 3..	I	—	—	—	—	4	—	—	2	—	—	—	—	—	—
) + 4..	I	—	—	—	—	2	—	—	—	—	—	—	—	—	—
) + 5..	—	—	—	—	—	2	I	—	—	—	—	—	—	—	—
) + 6..	—	—	—	I	—	I	—	—	—	—	—	—	—	—	—
) + 7..	—	I	—	—	—	—	—	—	—	—	—	—	—	—	—
5) + 0..	7	11	I	I	4	13	4	3	7	4	2	I	15	26	13
5) + 1..	6	7	3	8	7	9	6	5	8	4	11	2	31	20	7
5) + 2..	—	—	—	2	—	3	I	2	I	—	—	—	—	—	—
5) + 3..	—	—	—	—	—	I	—	2	I	—	—	—	—	—	—
5) + 0..	39	16	—	107	32	7	160	24	I	92	20	—	134	52	—
5) + 1..	—	I	—	12	6	2	10	2	I	5	2	—	25	6	I
5) + 2..	—	—	—	2	2	—	—	I	—	—	—	—	—	—	—
5) + 4..	—	—	—	—	I	—	—	—	—	—	—	—	—	—	—
7) + 0..	I	—	—	7	2	I	I	I	—	—	—	—	5	I	I
7) + 1..	—	—	—	—	—	—	—	I	—	—	—	—	4	—	—
7) + 2..	—	—	—	I	—	—	—	—	—	—	—	—	—	—	—
8) + 0..	—	—	—	I	—	—	—	—	—	—	—	—	I	—	—
8) + 1..	—	—	—	—	—	—	—	—	I	—	—	—	I	—	—
	56	43	99	142	57	199	183	43	226	106	42	305	221	114	420

Table 2 shows the average number of primary double bundles, intercalary bundles, and total bundles in the three types of seedlings, and gives the differences and probable errors of differences in the means upon which we must depend for conclusions.

The entries in the first section of this table show that the average number of primary double bundles is relatively lower in the hemitrimerous than in the trimerous seedlings. It is also relatively higher than the number in the dimerous seedlings. The differences, while small, may reasonably be considered significant in comparison with their probable errors. The differences between the hemitrimerous and the dimerous class are much larger than those between the hemitrimerous and the trimerous.

Turning to the statistical constants for intercalary bundles set forth in the second section of table 2, we note that in four of the five cases the hemitrimerous seedlings have a larger number of intercalary bundles than the trimerous seedlings. These differences are small, but may be significant. In the one case in which the hemitrimerous seedlings have a smaller number of intercalary bundles than the trimerous plantlets the difference is only − 0.01 ± 0.04. In two of the cases the hemitrimerous show a larger number of intercalary bundles than the dimerous seedlings, but in three lines the reverse is true. The differences are in general not so large in comparison with their probable errors as in the case of the comparison for number of primary double bundles.

TABLE 2. *Mean number of bundles at base of hypocotyl*

	f	Primary Double Bundles	Intercalary Bundles	Total Bundles
Line 29				
3–3	56	5.68 ± .05	.27 ± .07	5.95 ± .04
3–2	43	5.21 ± .08	.53 ± .12	5.74 ± .10
2–2	99	4.04 ± .01	.16 ± .03	4.20 ± .03
(3–2)–(3–3)		− 0.47 ± .09	+ .26 ± .14	− 0.21 ± .11
(3–2)–(2–2)		+ 1.17 ± .08	+ .37 ± .12	+ 1.54 ± .10
Line 75				
3–3	142	5.98 ± .02	.25 ± .04	6.23 ± .03
3–2	57	5.74 ± .05	.44 ± .07	6.18 ± .07
2–2	199	4.24 ± .03	.62 ± .05	4.85 ± .05
(3–2)–(3–3)		− 0.24 ± .05	+ .19 ± .08	− 0.05 ± .08
(3–2)–(2–2)		+ 1.50 ± .06	− .18 ± .09	+ 1.33 ± .09
Line 98				
3–3	183	5.93 ± .01	.13 ± .02	6.06 ± .02
3–2	43	5.67 ± .07	.53 ± .09	6.21 ± .08
2–2	226	4.11 ± .02	.62 ± .03	4.73 ± .04
(3–2)–(3–3)		− 0.26 ± .07	+ .40 ± .09	+ 0.15 ± .08
(3–2)–(2–2)		+ 1.56 ± .07	− .09 ± .09	+ 1.48 ± .09
Line 139				
3–3	106	5.91 ± .02	.09 ± .02	6.00 ± .02
3–2	42	5.36 ± .08	.43 ± .05	5.79 ± .06
2–2	305	4.01 ± .00	.13 ± .02	4.14 ± .02
(3–2)–(3–3)		− 0.55 ± .08	+ .34 ± .05	− 0.21 ± .06
(3–2)–(2–2)		+ 1.35 ± .08	+ .30 ± .05	+ 1.65 ± .06
Line 143				
3–3	221	5.81 ± .03	.29 ± .02	6.10 ± .03
3–2	114	5.45 ± .04	.28 ± .03	5.73 ± .04
2–2	420	4.06 ± .01	.29 ± .02	4.35 ± .02
(3–2)–(3–3)		− 0.36 ± .05	− .01 ± .04	− 0.37 ± .05
(3–2)–(2–2)		+ 1.39 ± .04	− .01 ± .04	+ 1.38 ± .04

We cannot, therefore, assert on the basis of the data now in hand whether dimerous, hemitrimerous, and trimerous seedlings differ in the number of intercalary bundles at the base of the hypocotyl. In so far as it goes the evidence *suggests* that the hemitrimerous seedlings have a larger number of intercalary bundles than the trimerous but a smaller number than the dimerous plantlets.

The means for total number of bundles (primary double bundles plus intercalary bundles) at the base of the hypocotyl set forth in the third section of table 2 show that in four of the five cases the mean number of bundles is lower in the hemitrimerous than in the trimerous seedlings. The differences are, however, very slight indeed and cannot in general be considered significant in comparison with their probable errors. The differences between the hemitrimerous and dimerous seedlings on the other hand are rather large and in every case are unquestionably significant.

Summarizing these results, we note that the hemitrimerous seedlings are conspicuously differentiated from the dimerous seedlings in the number of primary double bundles and in the total number of bundles. They are less conspicuously differentiated, if at all, in number of intercalary bundles. They are unquestionably differentiated from the trimerous seedlings by

heir lower number of primary double bundles and possibly by a higher umber of intercalary bundles. They cannot be said to differ from the rimerous seedlings in the total number of bundles at the base of the ypocotyl.

Central Region of Hypocotyl

For the number of bundles in the central region of the hypocotyl we ave the fundamental frequency distributions given in table 3. Considering he mean number of bundles in table 4, it appears that the number of bundles ı the central region of the hypocotyl of hemitrimerous plants is slightly ›wer than that found in trimerous seedlings in four of the six lines available. 'he differences are, however, small and would not for the most part be onsidered significant in comparison with their probable errors. The undle number of hemitrimerous plants is in every case distinctly higher han that of dimerous plants at this level, and these differences are conpicuous and unquestionably significant. Thus in hypocotyledonary strucure the hemitrimerous seedling is very close indeed to the trimerous but erhaps shows a slight deficiency in bundle number.

This result is not surprising in view of the fact that so far as the cotyedonary node and lower portions of the axis are concerned the external orm of hemitrimerous and trimerous seedlings is essentially identical.

Central Region of Epicotyl

If a differentiation between the hemitrimerous and trimerous seedlings ›btains anywhere, one would expect to find it in the epicotyledonary region,

TABLE 3. *Distribution of number of bundles in central region of hypocotyl*

	8	9	10	11	12	13	14	15	16	17	18	20	24	Total
ine 29														
3-3.....	—	—	1	6	41	1	3	3	—	—	—	—	1	56
3-2.....	—	2	6	8	21	2	2	—	1	1	—	—	—	43
2-2.....	67	21	9	2	—	—	—	—	—	—	—	—	—	99
ine 75														
3-3.....	1	3	5	36	292	40	29	5	1	4	—	—	—	416
3-2.....	—	2	3	13	51	16	14	2	2	—	—	—	—	103
2-2.....	177	131	103	46	31	14	11	1	1	3	1	—	—	519
ine 93														
3-3.....	—	—	8	32	382	82	38	12	1	—	1	1	—	557
3-2.....	—	—	4	6	17	8	7	1	—	—	—	—	—	43
2-2.....	36	93	170	107	96	40	18	1	—	—	2	—	—	563
ine 98														
3-3.....	—	1	6	12	297	21	8	—	—	—	—	—	—	345
3-2.....	—	—	3	7	25	3	3	1	1	—	—	—	—	43
2-2.....	125	126	83	37	11	5	—	—	—	1	—	—	—	388
ine 139														
3-3.....	—	—	4	8	84	6	3	1	—	—	—	—	—	106
3-2.....	1	3	4	11	19	3	1	—	—	—	—	—	—	42
2-2.....	269	23	7	4	1	1	—	—	—	—	—	—	—	305
ine 143														
3-3.....	2	1	11	14	136	21	25	6	3	1	1	—	—	221
3-2.....	1	2	17	19	54	4	8	5	4	—	—	—	—	114
2-2.....	263	83	47	17	4	3	1	2	—	—	—	—	—	420

since the sole superficial difference between the two types of seedlings is found at the primordial node. The frequency distributions in table 5 show that the nodal number of bundles is in general lower in the hemitrimerous than in the trimerous seedlings. It also indicates that they are higher in the hemitrimerous than in the dimerous seedlings. The averages and their probable errors in the second section of table 4 show that in each of

TABLE 4. *Mean number of bundles in central regions of internodes*

Line	f	Central Region of Hypocotyl	Central Region of Epicotyl	Line	f	Central Region of Hypocotyl	Central Region of Epicotyl
Line 29				Line 98			
3-3.....	56	12.36±.16	14.75±.18	3-3....	345	12.03±.02	14.89±.04
3-2.....	43	11.74±.16	12.93±.17	3-2....	43	12.07±.12	13.72±.13
2-2.....	99	8.45±.05	12.05±.02	2-2....	388	9.24±.04	12.12±.01
(3-2)-(3-3).		− .62±.23	− 1.82±.25	(3-2)-(3-3)		+ .04±.12	− 1.17±.14
(3-2)-(2-2).		+ 3.29±.17	+ .88±.17	(3-2)-(2-2)		+ 2.83±.13	+ 1.60±.13
Line 75				Line 139			
3-3.....	416	12.19±.03	15.47±.05	3-3....	106	11.99±.05	15.24±.08
3-2.....	103	12.32±.08	13.85±.10	3-2....	42	11.36±.12	13.93±.18
2-2.....	519	9.52±.05	12.26±.02	2-2....	305	8.19±.02	12.10±.01
(3-2)-(3-3).		+ .13±.09	− 1.62±.11	(3-2)-(3-3)		− .63±.13	− 1.31±.20
(3-2)-(2-2).		+ 2.80±.09	+ 1.59±.10	(3-2)-(2-2)		+ 3.17±.12	+ 1.83±.18
Line 93				Line 143			
3-3.....	557	12.29±.03	15.65±.04	3-3....	221	12.29±.06	16.10±.08
3-2.....	43	12.26±.13	14.84±.18	3-2....	114	11.89±.10	13.68±.09
2-2.....	563	10.62±.04	12.19±.02	2 -2....	420	8.66±.04	12.36±.02
(3-2)-(3-3).		− .03±.13	− .81±.18	(3-2)-(3-3)		− .40±.12	− 2.42±.12
(3-2)-(2-2).		+ 1.64±.14	+ 2.65±.18	(3-2)-(2-2)		+ 3.23±.11	+ 1.32±.09

TABLE 5. *Distribution of number of bundles in central region of epicotyl*

	10	11	12	13	14	15	16	17	18	19	20	21	22	Total
Line 29														
3-3.....	1	—	3	11	13	15	4	2	4	1	1	1	—	56
3-2.....	2	3	19	4	6	6	2	1	—	—	—	—	—	43
2-2.....	—	—	97	—	1	1	—	—	—	—	—	—	—	99
Line 75														
3-3....	—	—	3	16	63	164	93	41	27	4	4	1	—	416
3-2....	—	3	16	28	27	14	9	4	1	—	1	—	—	103
2-2.....	1	4	422	58	21	10	3	—	—	—	—	—	—	519
Line 93														
3-3.....	—	—	5	18	47	236	129	56	51	10	4	—	1	557
3-2.....	—	—	5	4	9	13	5	3	2	2	—	—	—	43
2-2.....	1	6	483	42	20	10	1	—	—	—	—	—	—	563
Line 98														
3-3.....	—	—	8	24	69	176	49	9	7	1	1	1	—	345
3-2....	—	1	6	12	13	8	2	1	—	—	—	—	—	43
2-2.....	—	—	352	27	7	1	1	—	—	—	—	—	—	388
Line 139														
3-3.....	—	—	—	8	21	38	24	9	4	2	—	—	—	106
3-2....	—	2	6	12	8	6	5	1	2	—	—	—	—	42
2-2....	—	—	278	23	4	—	—	—	—	—	—	—	—	305
Line 143														
3-3....	—	—	5	9	19	54	49	37	31	9	6	2	—	221
3-2.....	—	—	31	19	37	17	6	2	—	2	—	—	—	114
2-2.....	—	—	318	66	27	5	4	—	—	—	—	—	—	420

the six lines the average number of bundles in the epicotyl is significantly lower in the hemitrimerous than in the trimerous seedlings, and (probably) significantly higher in the hemitrimerous than in the dimerous seedlings.

In epicotyledonary structure the hemitrimerous seedlings occupy as a matter of fact almost exactly an intermediate position between the dimerous and the trimerous types.

·Summary

The purpose of this paper is a comparison of the gross vascular anatomy of hemitrimerous seedlings of *Phaseolus vulgaris* with those which are trimerous and those which are dimerous.   By dimerous seedlings we understand those with *two* cotyledons and *two* primordial leaves, by trimerous seedlings those with *three* cotyledons and *three* primordial leaves, and by hemitrimerous seedlings those with *three* cotyledons and *two* primordial leaves.   The hemitrimerous is, therefore, intermediate in external form between the dimerous and the trimerous seedling.   In the internal structure of the axis at the transition zone, which here occurs at the base of the hypocotyl, the hemitrimerous seedling is clearly differentiated from the trimerous type by a slightly smaller average number of primary double bundles, and possibly by a slightly larger number of intercalary bundles. The total number of bundles in the basal region of the axis of hemitrimerous seedlings is not sensibly different in hemitrimerous and trimerous plantlets. The hemitrimerous are conspicuously differentiated from the dimerous seedlings by a larger number of primary double bundles and a larger total number of bundles.   On the basis of the data available they cannot be asserted to differ significantly from the dimerous plants in the number of intercalary bundles.

In the central region of the hypocotyl, the vascular anatomy of the hemitrimerous seedling conspicuously exceeds that of the dimerous in bundle number but agrees very closely indeed with that of the trimerous plantlet, although it may have a slightly lower average number of bundles.

In the central region of the epicotyl the mean number of bundles in the hemitrimerous seedling is, roughly speaking, intermediate between that of the trimerous and that of the dimerous types.

Recapitulating, it appears that in internal structure the hypocotyl of the hemitrimerous seedling is practically identical with that of the trimerous seedling with which it has in common a whorl of three cotyledons.   The epicotyledonary internode in the hemitrimerous seedling, limited by a trimerous cotyledonary and a dimerous primordial node, is intermediate in anatomy between the trimerous type with three cotyledons and three primordial leaves and the dimerous type with two cotyledons and two primordial leaves.

THE INTERRELATIONSHIP OF THE NUMBER OF THE TWO TYPES OF VASCULAR BUNDLES IN THE TRANSITION ZONE OF THE AXIS OF PHASEOLUS VULGARIS

J. Arthur Harris, Edmund W. Sinnott, John Y. Pennypacker, and G. B. Durham

(Received for publication January 17, 1921)

Introductory

In papers[1] on the anatomy of dimerous[2] and trimerous and of hemi-trimerous seedlings we have shown that *Phaseolus vulgaris* is characterized by a structure of the vascular system at the base of the hypocotyl which is rather infrequent in seedling anatomy in general. This is the presence of a variable number of accessory bundles which usually lack protoxylem elements. These are the "Zwischenstränge" of Dodel. We have elsewhere called them intercalary bundles. They may make their appearance in the upper part of the root or in the lower region of the hypocotyl, some rising blindly below and others originating by division of a primary double bundle. These intercalary strands may be distinguished from the other bundles with perfect certainty because of their position and of the absence within them of any protoxylem elements.

In another place[3] we have dealt with the correlations between the number of bundles at different levels in the seedling, that is, the relationship between the vascular system at the base of the hypocotyl and that in the central region of the hypocotyl and epicotyl, and between the bundle system in the hypocotyl and that in the epicotyl. Our present problem is to consider the interrelationships between the two types of bundles present in the hypocotyl just above the region of transition from root to stem structures, and between each of these types and the total number of bundles in this zone.

[1] Harris, J. Arthur, Sinnott, E. W., Pennypacker, J. Y., and Durham, G. B. The vascular anatomy of dimerous and trimerous seedlings of *Phaseolus vulgaris*. Amer. Jour. Bot. 8: 63–102. 1921. The vascular anatomy of hemitrimerous seedlings of *Phaseolus vulgaris*. Amer. Jour. Bot. 8: 375–381. 1921.

[2] Dimerous seedlings have 2 cotyledons and 2 primordial leaves; trimerous seedlings have 3 cotyledons and 3 primordial leaves; and hemitrimerous seedlings have 3 cotyledons and 2 primary leaves.

[3] Harris, J. Arthur, Sinnott, E. W., Pennypacker, J. Y., and Durham, G. B. Correlations between anatomical characters in the seedling of *Phaseolus vulgaris*. Amer. Jour. Bot. 8: 339–365. 1921.

Lack of space precludes the publication of the 30 individual correlation tables upon which the coefficients discussed in this section are based. These may, however, be easily formed from the schedules showing the formula for the basal bundles in other papers of this series.[4]

TABLE I. *Correlation between Number of Primary Double Bundles and Number of Intercalary Bundles at Base of Hypocotyl*

Line	Trimerous			Dimerous			Difference	Diff. $\overline{E_{\text{diff.}}}$
	N	r	$\dfrac{r}{E_r}$	N	r	$\dfrac{r}{E_r}$		
75	142	$-.5004\pm.0424$	11.8	142	$-.1177\pm.0558$	2.11	$-.3827\pm.0700$	5.46
93	155	$-.6155\pm.0337$	18.3	155	$-.1449\pm.0530$	2.73	$-.4706\pm.0624$	7.54
98	183	$-.6515\pm.0286$	22.8	183	$+.0643\pm.0496$	1.30	$-.7158\pm.0574$	12.4
139	106	$-.5053\pm.0488$	10.4	305	$+.1364\pm.0379$	3.60	$-.6417\pm.0618$	6.0
143	221	$-.3184\pm.0408$	7.8	420	$+.0338\pm.0329$	1.03	$-.3522\pm.0530$	5.4

ANALYSIS OF DATA

1. *Relationship between Number of Primary Double Bundles and Number of Intercalary Bundles.* We shall first consider the relationship between the number of primary double bundles and the number of intercalary bundles at the base of the hypocotyl in dimerous and trimerous plants

The correlation coefficients for the five lines appear in table I. For the trimerous plants of all five lines the correlations are negative in sign *i.e.*, the number of intercalary bundles is greater in plants which have a smaller number of primary double bundles, and *vice versa*. For dimerous plants three of the five lines show a slightly negative coefficient, but two show a low positive correlation. The constants indicate that the corre lations for the trimerous plants are much higher numerically than those for the dimerous plants. Those for the trimerous are of the order $-$. to $-.6$ while those for dimerous plants are sensibly zero, averaging $+.005$ The correlations for the trimerous plants are in all cases several times a large as their probable errors, while those for the dimerous plants could hardly be regarded as statistically significant if only one of the lines were available. The differences, taken with regard to sign, between the corre lations for the dimerous and trimerous plants are in each case significant in comparison with their probable errors.

Expressing these results in terms of regression we have the following equations:

---

[4] The entries to be selected from the published tables can be determined from the value of N. In lines in which true siblings were available (75, 93, and 98) only siblings have been used, even though additional sections of one or the other type were available. In the two lines in which random samples of seed were used for the production of the dimerous and trimerous seedlings, the largest possible number of individuals available in the tables of the papers cited was employed for the constants here discussed.

	Dimerous	Trimerous

Line 75: $P = 4.255 - 0.058\ I$ $P = 6.059 - 0.318\ I$
$\qquad\qquad I = 1.641 - 0.239\ P$ $I = 4.968 - 0.789\ P$

Line 93: $P = 4.607 - 0.110\ I$ $P = 5.992 - 0.448\ I$
$\qquad\qquad I = 1.398 - 0.127\ P$ $I = 5.200 - 0.846\ P$

Line 98: $P = 4.099 + 0.035\ I$ $P = 5.984 - 0.392\ I$
$\qquad\qquad I = 0.114 + 0.117\ P$ $I = 6.559 - 1.084\ P$

Line 139: $P = 4.005 + 0.034\ I$ $P = 5.958 - 0.558\ I$
$\qquad\qquad I = -2.038 + 0.541\ P$ $I = 2.795 - 0.457\ P$

Line 143: $P = 4.060 + 0.019\ I$ $P = 5.924 - 0.408\ I$
$\qquad\qquad I = 0.047 + 0.059\ P$ $I = 1.732 - 0.249\ P$

The mean number of intercalary bundles associated with given numbers of primary double bundles and the theoretical means as given by the regression straight lines are shown on diagram 1.

For the normal plants of lines 75, 93, and 139 the agreement between the observed means and the regression line is very satisfactory. In line 98 a single seedling with 8 primary double bundles and 4 intercalary bundles gives a positive sign to the correlation and makes the agreement of theoretical and empirical means very poor indeed. In lines 139 and 143 the correlation is also positive. It must be noted that we are dealing here with a very narrow range of both primary double bundles and intercalary bundles, and with very small frequencies in some of the classes.

For the abnormal plants the agreement of empirical means and theoretical lines is apparently very poor indeed. This is perhaps largely attributable to two facts:

(a) The frequencies of primary double bundles are, practically speaking, concentrated in two classes, 5 and 6 bundles. From 93 to 99 percent of the seedlings fall in these two classes. As a result of this condition, the obtaining of trustworthy averages for the extreme classes of primary double bundles is, practically speaking, impossible.

(b) The influence of the two principal classes (5 and 6) of primary double bundles is such as to throw the theoretical mean number of intercalary bundles for higher classes of primary double bundles on the negative side of 0 in four of the five cases.

As a consequence, the actual mean number of intercalary bundles must lie above the line in all cases in which more than 6 primary double bundles are formed.

Whether these irregularities represent a significant deviation from linearity can be determined only when far larger series of data are available.

While the primary double bundles must probably be regarded as more fundamental structures than the intercalary bundles, it seems of interest to determine the mean number of primary double bundles associated with each number of intercalary bundles.

The lines for the regression of number of primary double bundles on number of intercalary bundles are represented with the empirical means on diagram 2. These figures show that, with the exception of the normal plants of lines 98, 139, and 143, the number of primary double bundles decreases slightly as the number of intercalary bundles increases. The rate of decrease is somewhat greater in the abnormal than in the normal plants.
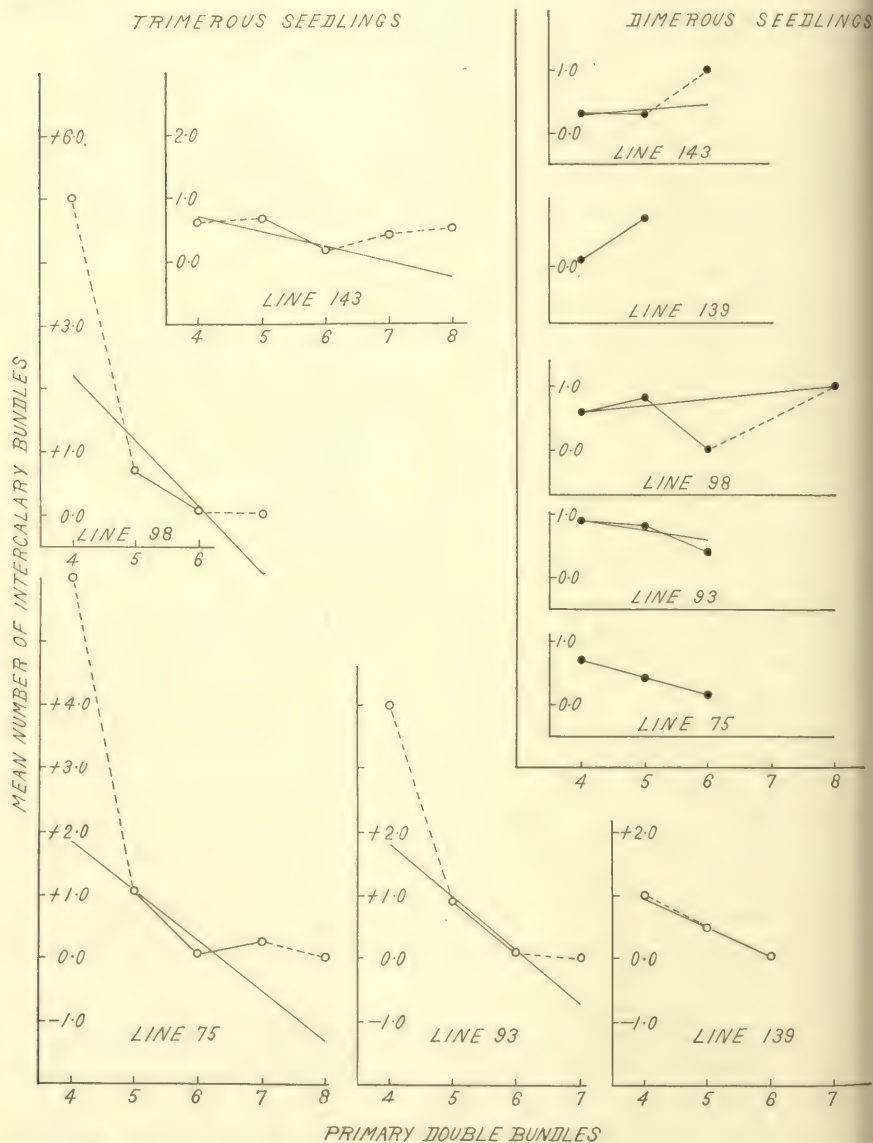

DIAGRAM 1.   Regression of number of intercalary bundles on number of primary double bundles, at base of hypocotyl.

It is suggestive to note that the negative correlation between number of primary double bundles and number of intercalary bundles demonstrated here within seedlings of one class with regard to external structure is also evident when we pass from a type of seedling with a smaller to one with a higher number of primary double bundles.   It has been shown in an earlier paper that (a) the number of trimerous seedlings having intercalary bundles is generally smaller than the number of dimerous seedlings with these accessory structures, and that (b) the average number of intercalary bundles is generally smaller in trimerous than in dimerous seedlings.

DIAGRAM 2.   Regression of number of primary double bundles on number of intercalary bundles, at base of hypocotyl.

2. *Relationship between the Total Number of Bundles and the Number of Bundles of the Two Types.* We may now inquire to what extent the variation in the total number of bundles depends upon the primary double bundles and to what extent upon the number of intercalary bundles.   As a first step we determine the correlation between the total number of bundles and the number of primary double bundles, and between the total number of bundles and the number of intercalary bundles.   These results are set forth in table 2.

We note that the correlations between the total number of bundles and the number of intercalary bundles are in all cases high, ranging from $+.42$

to $+.80$ in the trimerous and from $+.76$ to $+.97$ in the dimerous plants. The correlations for the dimerous plants are in all five cases slightly higher than those for the trimerous plants.

TABLE 2. *Comparison of Correlation between Total Bundles and Primary Double Bundles, r_{bp}, and between Total Bundles and Intercalary Bundles, r_{bi}, at Base of Hypocotyl*

	Trimerous			Dimerous			Difference	Diff. E_{diff}
	N	r	$\dfrac{r}{E_r}$	N	r	$\dfrac{r}{E_r}$		
Line 75								
r_{bi}.....	142	$+.7788\pm.0222$	35.1	142	$+.8872\pm.0120$	73.9	$-.1084\pm.0245$	4.42
r_{bp}.....	142	$+.1532\pm.0552$	2.77	142	$+.3536\pm.0495$	7.14	$-.2004\pm.0741$	2.70
$r_{bi}-r_{bp}$.		$+.6256\pm.0591$	10.6		$+.5336\pm.0510$	10.5		
Line 93								
r_{bi}.....	155	$+.6934\pm.0281$	24.7	155	$+.7628\pm.0226$	33.8	$-.0694\pm.0361$	1.92
r_{bp}.....	155	$+.1409\pm.0531$	2.65	155	$+.5292\pm.0390$	13.6	$-.3883\pm.0655$	5.92
$r_{bi}-r_{bp}$.		$+.5525\pm.0600$	9.21		$+.2336\pm.0447$	5.23		
Line 98								
r_{bi}.....	183	$+.8001\pm.0179$	44.6	183	$+.8833\pm.0109$	81.0	$-.0832\pm.0200$	4.16
r_{bp}.....	183	$-.0664\pm.0496$	1.34	183	$+.5245\pm.0361$	14.5	$-.5909\pm.0616$	9.59
$r_{bi}-r_{bp}$.		$+.8665\pm.0529$	16.4		$+.3588\pm.0374$	9.59		
Line 139								
r_{bi}.....	106	$+.4203\pm.0539$	7.79	305	$+.9721\pm.0021$	457.4	$-.5518\pm.0539$	10.2
r_{bp}.....	106	$+.5707\pm.0422$	12.9	305	$+.3649\pm.0335$	10.9	$+.2058\pm.0555$	3.71
$r_{bi}-r_{bp}$.		$-.1504\pm.0697$	2.16		$+.6072\pm.0336$	18.1		
Line 143								
r_{bi}.....	221	$+.4382\pm.0367$	11.9	420	$+.8715\pm.0079$	110.2	$-.4333\pm.0375$	11.55
r_{bp}.....	221	$+.7126\pm.0223$	31.9	420	$+.5196\pm.0240$	21.6	$+.1930\pm.0328$	5.88
$r_{bi}-r_{bp}$.		$-.2744\pm.0429$	6.40		$+.3519\pm.0253$	13.9		

The correlation between the total number of bundles and the number of primary double bundles is in general much lower. In line 98 the coefficient actually has the negative sign in the trimerous series. The differences between the correlation coefficients for total bundles and intercalary bundles, and for total bundles and primary double bundles, range from $-.27$ to $+.87$ in the trimerous plants and from $+.23$ to $+.61$ in the dimerous plants.

It is clear that the two types of plants differ rather fundamentally in this correlation. The correlation between the total bundles and the primary double bundles is very low in the trimerous plants. It is a much more substantial value in the dimerous plants.

Pursuing this point one step farther, we may determine by a special formula the relationship between the total number of bundles and the deviation of the number of intercalary bundles from the number which would be expected if the number of primary double bundles and intercalary bundles were in proportion to the total number of bundles formed.

Determining the correlation between the total number of bundles, b,

nd the deviation of the number of intercalary bundles, i, from their proba-
ble value by the formula[5]

$$r_{bz} = \frac{r_{bi} - r_b/r_i}{\sqrt{1 - r_{bi}^2 + (r_{bi} - r_b/r_i)^2}}$$

$$\text{where } z = i - \frac{\bar{i}}{\bar{b}}b,$$

we have the values given in table 3.

TABLE 3. *Correlation between Total Bundles at Base of Hypocotyl and Deviation of Number of Intercalary Bundles from Their Probable Number*

Line	Trimerous			Dimerous			Difference	Diff. $E_{\text{diff.}}$
	N	r	$\frac{r}{E_r}$	N	r	$\frac{r}{E_r}$		
75	142	.7643±.0235	32.5	142	.8513±.0156	54.6	−.0870±.0283	3.07
93	155	.6787±.0292	23.2	155	.6693±.0299	22.4	+.0094±.0412	0.22
98	183	.7944±.0184	43.2	183	.8433±.0144	58.6	−.0489±.0224	2.18
139	106	.4066±.0546	7.45	305	.9701±.0023	421.8	−.5636±0.540	10.3
143	221	.3841±.0386	9.95	420	.8510±.0090	94.6	−.4669±03.96	11.8

The coefficients are positive and high, and very consistent for the two
types of seedlings. They show that within one morphological type of
seedling[6] an increase in the total number of bundles is primarily due to the
formation of intercalary bundles, rather than to variation in the number of
primary double bundles, although both types of bundles contribute to the
end result.

SUMMARY

An investigation of the interrelationship of the numbers of primary
double bundles, intercalary bundles, and total bundles (primary double
bundles plus intercalary bundles) at the base of the hypocotyl in dimerous
and trimerous seedlings of *Phaseolus vulgaris* leads to the following results:

1. In the trimerous seedlings there is a negative correlation of about
medium value ($r = -.5 \pm$) between the number of primary double bundles
and the number of intercalary bundles. Thus the number of intercalary
bundles is smaller in seedlings with larger numbers of primary double
bundles and *vice versa*. In dimerous seedlings the correlation is perhaps
also negative in sign, but practically zero numerically.

[5] Harris, J. Arthur. The correlation between a variable and the deviation of a dependent variable from its probable value. Biometrika 6: 438–443. 1909; also, Further illustrations of the applicability of a coefficient measuring the correlation between a variable and the deviation of a dependent variable from its probable value. Genetics 3: 328–352. 1918.

[6] The differentiation of trimerous and dimerous seedlings has been shown to be due primarily to an increase in the number of primary double bundles.

This result for seedlings of the same morphological type is suggestive in its relation to the results of a comparison of seedlings which are externally dimerous and trimerous, since in general trimerous seedlings show an increase in number of primary double bundles but a decrease in number of intercalary bundles as compared with dimerous seedlings. As a result of this numerical compensation, most conspicuously evident in the trimerous seedlings, the total number of bundles shows a lower variability than it would if the numbers of the two types of bundles were quite independent.

2. The correlation between the total number of bundles (primary double bundles plus intercalary bundles) and the number of intercalary bundles is high. The coefficients for the dimerous seedlings are somewhat higher than those for the trimerous seedlings. The correlation between the total number of bundles and the number of primary double bundles is generally much lower. The correlation between the total number of bundles and the deviation of the number of intercalary bundles from that which would be expected if they occurred in the same proportionate frequency throughout the entire range of total bundle number is positive in sign and substantial in magnitude. In both types of seedlings variation in the number of intercalary bundles is therefore an important factor in determining variation in the total number of bundles at the base of the hypocotyl.

THE VASCULAR ANATOMY OF DIMEROUS AND TRIMEROUS SEEDLINGS OF PHASEOLUS VULGARIS

J. ARTHUR HARRIS, EDMUND W. SINNOTT, JOHN Y. PENNYPACKER, AND G. B. DURHAM

INTRODUCTORY

The great majority of investigations dealing with the anatomy of plants have been purely descriptive in character. As a result of observation, the typical or average condition of plant structures has been recorded in terms which are general and often indefinite. Comparatively few morphological papers deal with the problem of the variation of the structures under consideration, treat of their correlations with one another, or even present the detailed measurements which might serve for the solution of such fundamental morphological problems.

The older comparative morphology is indispensable. It provides a general knowledge of plant structures and serves as a basis for the classification of the vegetable kingdom. The recognition that description must be supplemented by the results of experimentation has, however, led to the establishment of the newer special science of experimental morphology. The time has come to extend still further our study of plant form by calling to the service of vegetable morphology the methods of measurement and mathematical analysis. These methods are particularly useful in an attack upon the fundamental problems of morphogenesis. It is by measuring exactly the various plant structures during their successive stages of development, in terms of size or number; by determining their relative variability in different organs or regions of the plant, or under varying external conditions, and by discovering such correlations as exist, both among the structures themselves and between them and their progenitors and their environment, that we shall be able to build up a body of fact on which morphogenetic theory may rest.

The present paper gives a portion of the results of a biometric analysis of a comparatively simple morphological problem, that of the gross vascular anatomy of certain normal and abnormal bean seedlings. Our purpose has been:

1. A study of the vascular anatomy of normal and of abnormal seedlings from the point of view of descriptive morphology—a preliminary which we believe to be essential to a sound interpretation of any statistical results.

2. A statistical study of the number and variation of the vascular elements in different regions of the seedling.

3. An investigation of the correlations between these internal characters (such as those which exist between bundle number in different regions of the seedling) and between the internal characters and external features of the plant.

The results of the first and second phases of the investigation are set forth in the present paper; the third is reserved for a later publication.

MATERIALS AND METHODS

A priori considerations seemed to indicate that a promising line of attack upon the general field of quantitative plant morphology lay in the investigation of vascular bundle number. Such an investigation should be on a scale sufficiently large to make possible the determination of trustworthy biometric constants, and should have as its subject a plant organ of relatively simple but variable structure. Because of the ease with which they can be grown in quantity, their sharply marked external characteristics, their convenient size for histological work, and their relatively simple internal structure, seedlings of *Phaseolus vulgaris* furnish highly satisfactory material for a study of variation and correlation in vascular structures.

Among the many types of variant seedlings of the garden beans which may be secured by extensive plantings, two were selected for investigation: (*a*) normal (*dimerous*) seedlings, with two cotyledons and two primordial leaves, and (*b*) *trimerous* seedlings, with three cotyledons and three primordial leaves. For brevity in table headings the dimerous plants will sometimes be represented by "2–2" and the trimerous by "3–3," where the first figure gives the number of cotyledons and the second the number of primordial leaves.

Since one of the purposes of this work is to carry out a comparison of bundle number in normal and teratological seedlings, the selection of a satisfactory control series of normal plants is a matter of primary importance. It is essential that the seedlings of the types to be compared be selected in a manner to reduce to a minimum any external influences tending to bring about differences between them. It is clear that if the abnormal and the normal seedlings were taken from different series of parent plants, either genetic differences or environmental influences acting upon the parent plant might be effective in bringing about a differentiation in the characters of the seedling examined. A normal seedling from the same parent was, therefore, taken for comparison with each abnormal seedling[1] in each series in which the seed was derived from individual parent plants. Closer control of the influence of innate differences in the parents and of the possible influence of parental environment hardly seems practicable since the

---

[1] In the vast majority of the cases one abnormal seedling only was sectioned from a parent plant. When more than one abnormal seedling was available a control was taken for each. Naturally it is immaterial whether control *a* or *b* be compared with abnormal seedling *A* or *B*, since all are siblings.

pairs of abnormal and normal seedlings were, in three of the lines investigated, derived from the same parent plant.

Furthermore, care was taken that seedlings compared were grown under essentially identical conditions, in order to reduce to a minimum the environmental influences which might possibly tend to bring about differences between them. Seeds from individual plants were germinated in flats and harvested as soon as possible after they broke through the sand. Thus all seeds not only developed under the same parental environment but were germinated under sensibly identical conditions, were collected simultaneously, and were in consequence sectioned at essentially the same stage of maturity.

Because of the rapidity with which seedlings change and the great influence of temperature upon growth, it is difficult to standardize, or exactly to describe, the stage of development at which the seedlings were taken. Most of them were placed in alcohol before or very soon after the primordial leaves had unfolded. Thus a fairly uniform and early stage of development was secured.[2]

Free-hand sections were cut and mounted temporarily. When necessary, phloroglucin and hydrochloric acid were used to bring out the vascular bundles. The general vascular topography of the seedlings was studied, but the data for the statistical analysis of the seedling anatomy were derived from a careful count of the number of vascular bundles at various levels in the seedling. Because of a certain amount of variation in the number of bundles with position in the organ, counts were made in definite regions only—the basal region of the hypocotyl (just at the point of transition from "root structure" to "stem structure"); the median region of the hypocotyl; and the median region of the epicotyl. In three series counts were also made of the protoxylem poles in the upper portion of the primary root.

The number of data available for the several regions differs because of a change in the plan of the work. Sectioning and counting were begun by two of us at Cold Spring Harbor in the summer of 1917 and continued with the assistance of Miss Eunice Kinnear in the summer of 1918. This work was confined to the mid-regions of the hypocotyl and epicotyl. From a statistical study of these data it seemed desirable to have a further series of countings made independently by a specialist in vascular anatomy. The work was, therefore, continued at Storrs during 1918, 1919, and 1920. We are greatly indebted to Miss Flora Miller for assistance in this phase of the work. At Storrs, sections were made at the base of the hypocotyl as well as in the mid-region of hypocotyl and epicotyl. In three series, sections were made of the root as well.

The bundles vary considerably in size, the largest being well developed

---

[2] Some of the seedlings of line 143 were allowed to become a little older, but there is no evidence of change in bundle number with age.

and the smallest containing only one or two lignified xylem cells and a small patch of phloem. Some are even more reduced, consisting of a phloem patch alone. Any strand in which at least one well lignified xylem element could be made out was counted as a bundle. Some of the bundles are partially double in character, this condition being due either to partial fusion or to incipient division. Whenever such a strand was surrounded by one bundle sheath it was counted as one bundle; when the separation was so great that the bundle sheath itself showed signs of division, the strand was counted as two.

The seedlings were harvested at a stage when the vascular tissues of the first epicotyledonary internode were not completely lignified, and the number of bundles counted was therefore possibly less than the number which would finally be developed there.

None of these possible sources of error is believed to be great enough to affect the conclusions appreciably.

The Structure of the Seedling

In order to provide a sound basis for the understanding and interpretation of our later work, it is necessary to present a brief descriptive account of the structure of the seedlings.

The Normal (Dimerous) Seedling

The morphology of the seedling of Phaseolus has received the attention of several investigators, notably Dodel[3] and Compton.[4] Like most of the large seedlings of the Leguminosae it is normally tetrarch in fundamental plan; that is, there are four groups of protoxylem elements in the root. At a very early stage there is associated with each of these a group of metaxylem cells. It is these groups of metaxylem elements, throughout the whole seedling, which in the present paper are counted as "bundles," even though (as is sometimes the case) they are not associated with protoxylem clusters.

At the stage when these seedlings were harvested, cambial activity had hardly begun to show itself, so that these primary bundles remained distinct and easy to identify.

The condition in the upper part of the root of a normal seedling is shown in figure 1. The four bundles, two in the cotyledonary plane and two in the intercotyledonary plane, are more or less V-shaped (with the protoxylem group in an exarch position at the apex of the V) and tend to extend laterally. They surround a large pith. In passing up into the base of the hypocotyl, each of these bundles divides into two (fig. 2), and typical stem structure,

---

[3] Dodel. A. Der Übergang des Dicotyledonen-stengels in die Pfahl-wurzel. Pringsh. Jahrb. 8: 149–193. 1872.

[4] Compton, R. H. An investigation of the seedling structure in the Leguminosae. Jour. Linn. Soc. 41: 7–122. 1912.

vth the protoxylem in an endarch position, begins to be assumed.   Each
ṛir is subsequently referred to as a "primary double bundle."   Thus the
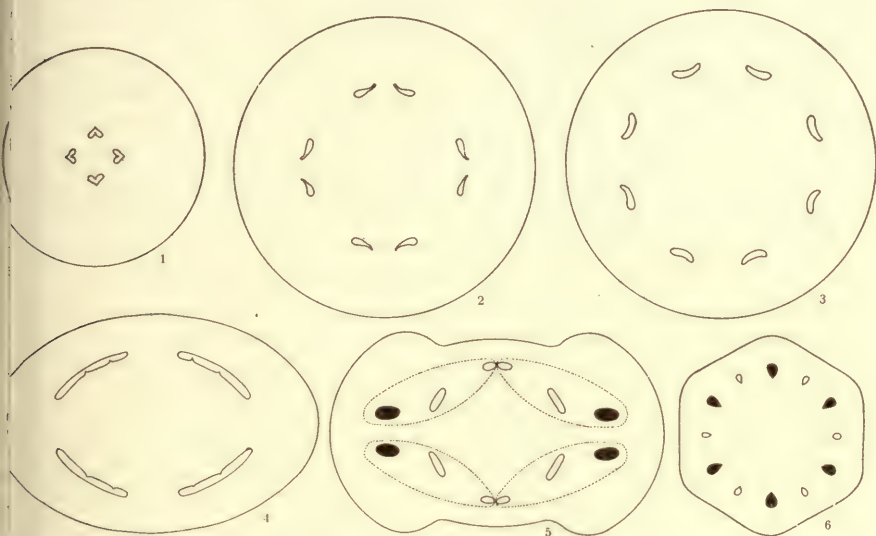l⸝el of transition from root structure to stem structure is low, being prac-

FIG. 1.   Dimerous seedling.   Transverse section through the root, showing its
t:rarch condition (four protoxylem poles). FIG. 2. Dimerous seedling.   Transverse
s:tion through the base of the hypocotyl showing the four primary double bundles, each
c which has been derived from one of the four root strands.   FIG. 3.   Dimerous seedling.
⸝ansverse section through the mid-region of the hypocotyl showing the normal eight-
ḷndled condition.   No intercalary bundles are figured.   FIG. 4.   Dimerous seedling.
⸝ansverse section just below the cotyledonary node.   The four bundles or bundle groups
ḷve originated by a more or less complete fusion of the adjacent members of each of the
c̣ginal pairs.   Each bundle, as shown by the two constrictions in it, is about to break up
i:o the three strands shown in figure 5.   FIG. 5.   Dimerous seedling.   Transverse section
t̠rough the cotyledonary node.   Each group of three strands which have arisen by a
ḷeaking up of the large bundles in figure 4 is here enclosed by a dotted line.   These three
s̠ands are a cotyledonary trace (solid black), an epicotyledonary bundle, and a small
ḷndle which will fuse with its adjacent neighbor to form another epicotyledonary bundle.
⸝G. 6.   Dimerous seedling.   Transverse section through the mid-region of the epicotyl
s̠owing the twelve bundles which have arisen by the splitting of the six original epicoty-
ḷlonary bundles.   The six strands which are to go off as traces to the two primordial
ḷves are solid black.

cally at the base of the hypocotyl.   The members of each of these four
ṛirs soon separate until the eight bundles are approximately equidistant
ḷg. 3), a condition which persists throughout the hypocotyl until the coty-
ḷdonary node is approached.

    In addition to these bundles, there are in a considerable percentage of
ḷe normal seedlings studied a variable number of accessory or intercalary
ṇundles, the "Zwischenstränge" of Dodel.   These may make their appear-

ance in the upper part of the root or in the lower region of the hypocotyl, some ending blindly below and others arising by division of the primary bundles. These intercalary bundles, which are not a very common feature of seedling anatomy in general, perhaps serve to increase the conductive capacity of the hypocotyl and may be associated with the large size of the seedling. They usually lack protoxylem elements.

At the cotyledonary node there is a rather complex anastomosis of the bundle system. The details of this vary somewhat, but its fundamental features are as follows: The two members of each of the two original pairs of bundles in the cotyledonary plane (that is, opposite the two points where the cotyledons will later arise) become widely separated, and each member fuses with the adjacent member of the intercotyledonary pair (fig. 4). Four large bundles or bundle aggregates are thus produced. Each breaks up immediately, usually into three parts. The lateral member of each group of three which is in the *cotyledonary* plane approaches the corresponding bundle of the next group of three, and these two strands become the cotyledonary traces and enter the base of the cotyledon. The lateral member of each group of three which is in the *intercotyledonary* plane approaches the corresponding bundle of the next group and fuses with it. The changes which are made and the resultant condition at this stage are shown in figure 5. Two strands (solid black) are here departing to each cotyledon, and six bundles are left as the basis for the vascular system of the epicotyl. The details of this nodal complex vary somewhat owing to the different levels at which fusion and separation of bundles take place, and to the presence of intercalary bundles. These intercalary bundles, as they approach the cotyledonary node, fuse with the others and are completely lost, exactly six epicotyledonary strands almost invariably emerging from the complex, quite regardless of the number of intercalary bundles which may have entered it from the hypocotyl. This fact we shall find to be of importance when we consider the statistical relationships of bundle number in hypocotyl and epicotyl.

Above the cotyledons, the six remaining bundles approach one another closing the cotyledonary gaps and forming a ring, the members of which almost immediately divide. The twelve bundles thus produced (fig. 6) persist throughout the first internode of the epicotyl.

At the first node of the epicotyl are inserted the two primordial leaves Phaseolus, like other Leguminosae which have been investigated, possesses a trilacunar node, the leaf being supplied by three traces, each of which causes a separate gap in the vascular ring.[5] The two primary leaves therefore remove six of the twelve bundles of the epicotyl (solid black in fig. 6). The six new bundles which appear just above the cotyledonary node are therefore, evidently downwardly extending leaf traces. These facts make

[5] Sinnott, E. W. The anatomy of the node as an aid in the classification of Angiosperms. Amer. Jour. Bot. **1**: 303–322. 1914.

understandable the almost invariably twelve-bundled condition of the first epicotyledonary internode.

The structure of the normal seedling thus corresponds to the type found by one of the writers[6] to be characteristic of a large number of Angiosperm families, in which the vascular supply to each cotyledon, consisting of two strands, leaves but one gap in the vascular ring; and in which the foliage leaf is trilacunar.

The Trimerous Seedling

The seedling with three cotyledons and three primordial leaves is built on a different plan from the normal one in that it is prevailingly hexarch, six

FIG. 7. Trimerous seedling. Transverse section through the root, showing its hexarch condition. FIG. 8. Trimerous seedling. Transverse section through the base of the hypocotyl, showing the six primary double bundles. FIG. 9. Trimerous seedling. Transverse section through the mid-region of the hypocotyl, showing the normal twelve-bundled condition. FIG. 10. Trimerous seedling. Transverse section just below the cotyledonary node. The six bundles or bundle groups correspond in origin and character to the four bundles of the dimerous seedling at this level. FIG. 11. Trimerous seedling. Transverse section through the cotyledonary node. Each group of three strands bounded by a dotted line corresponds in origin and character to a similar group at this level in the dimerous seedling. FIG. 12. Trimerous seedling. Transverse section through the mid-region of the epicotyl, showing the eighteen bundles which have arisen by the splitting of the nine original epicotyledonary bundles. The nine strands which are to go off as traces to the three primordial leaves are solid black.

[6] Sinnott, E. W.   Conservatism and variability in the seedling of dicotyledons.   Amer. Jour. Bot. 5: 120–130.   1918.

bundles occurring in the upper part of the root (fig. 7). This number is soon reduced to five and eventually to four, in passing down the root.

Passing upward into the hypocotyl, the six main strands (the primary double bundles) divide to produce twelve (figs. 8 and 9). Intercalary bundles are much less common than in the normal seedlings, appearing in only a small percentage of cases, and then being rarely more than one or two in number. At the node the same general procedure is followed as in the normal seedling, except, of course, that there are more bundles concerned. Bundles of adjacent pairs approach and fuse (fig. 10). Each of these bundles or bundle aggregates then divides, generally into three. Three cotyledons are each supplied with two bundles (solid black), and three sets of three bundles each—each formed by the fusion of two lateral bundles in the intercotyledonary plane—remain behind. The bundle changes and the final condition at the departure of the cotyledonary traces are shown in figure 11. The epicotyledonary ring which forms from the bundles which remain thus consists of nine strands instead of the normal six. Many of these divide at once, although the number is not usually doubled, as in normal seedlings, but varies from 12 to 18 or even more in the mid-region of the epicotyl (fig. 12). The bundles are much more crowded than in the normal seedlings, which may perhaps account for the failure of some of them to divide at once.

A study of the first epicotyledonary node shows that three strands are given off to each primary leaf, leaving from 6 to 9 in the stem.

It is therefore evident that within classes of seedlings which are uniform externally there are considerable anatomical variations and that the two classes investigated are profoundly differentiated in their anatomical organization.

Our next task is to subject the mass of data upon which these general conclusions are based to a statistical analysis with the object of bringing out otherwise undeterminable relationships.

BUNDLE NUMBER AND ITS VARIATION AT DIFFERENT LEVELS IN THE SEEDLINGS

From the statistical side we have two problems to consider.

The first is that of the relative numbers of bundles at different levels, *i.e.,* in the root, at the base of the hypocotyl, in the central region of the hypocotyl, and in the epicotyl of the same plant in both normal and abnormal plants, together with the variability in bundle number in different regions.

The second is that of the differences in bundle number, and in variation of bundle number, between normal and abnormal plants.

Since it is impossible to consider type and variation of bundle number at different levels without noting differences in the trimerous and dimerous forms upon which the observations were based, we shall devote this section primarily to a parallel discussion of both problems.

We shall consider in order the levels at which sections were made, beginning at the root.

1. *Root.* Roots were sectioned in the cases of lines 93, 139, and 143. The numbers of bundles[7] in the roots of normal and trimerous seedlings of these lines are shown in table 1.

TABLE 1

Primary Double Bundles	Line 93		Line 139		Line 143	
	Trimerous	Dimerous	Trimerous	Dimerous	Trimerous	Dimerous
3	—	—	2	—	4	—
4	31	132	15	149	37	219
5	87	20	53	1	113	2
6	34	—	36	—	66	—
7	—	—	—	—	1	—

The entries in this table show that most of the normal plants are tetrarch, although a small percentage are pentarch. In the trimerous seedlings the highest percentage are pentarch, but the remainder are distributed between tetrarch and hexarch with a few in more extreme classes. Sections made at progressively lower levels in the root show that the hexarch and pentarch conditions, in the trimerous seedlings, soon give way to tetrarch. This fact doubtless explains the relatively large number of non-hexarch cases

TABLE 2. *Vascular formula for base of hypocotyl of trimerous seedlings and their normal controls*

Base of hypocotyl	Line 75		Line 93		Line 96		Line 139		Line 143	
	Trimerous	Dimerous	Trimerous	Dimerous	Trimerous	Dimerous	Trimerous	Dimerous	Trimerous	Dimerous
4	—	69	—	34	—	97	—	138	2	150
4 + 1	—	30	—	37	—	43	1	9	3	55
4 + 2	—	10	—	13	—	23	—	—	—	4
4 + 3	—	4	—	5	—	2	—	—	—	—
4 + 4	—	2	1	1	—	—	—	—	—	—
4 + 5	—	2	—	—	1	—	—	—	—	—
4 + 6	1	—	—	—	—	—	—	—	—	—
5	1	13	5	22	4	6	4	1	15	5
5 + 1	8	4	10	18	6	8	4	2	31	5
5 + 2	2	1	3	9	1	1	—	—	—	—
5 + 3	—	1	—	1	—	1	—	—	—	—
6	107	5	120	10	160	1	92	—	134	—
6 + 1	12	1	11	3	10	—	5	—	25	1
6 + 2	2	—	1	2	—	—	—	—	—	—
7	7	—	4	—	1	—	—	—	5	1
7 + 1	—	—	—	—	—	—	—	—	4	—
7 + 2	1	—	—	—	—	—	—	—	—	—
8	1	—	—	—	—	—	—	—	1	—
8 + 1	—	—	—	—	—	1	—	—	1	—
	142	142	155	155	183	183	106	150	221	221

[7] Where the bundles were united in a ring, the number refers to number of protoxylem strands.

observed, for the zone within which the hexarch condition persists is narrow and its level is variable; and there is necessarily more or less variation in the level at which the sections are cut.

2. *Base of Hypocotyl.* In the series of sections of the base of the hypocotyl made at Storrs, the number of double vascular strands (each of which is derived from a primary root bundle and corresponds to a pole of the root) and the number of intercalary strands were recorded separately. There is no difficulty in distinguishing between these two categories of bundles, since the latter are almost invariably without protoxylem elements and are irregularly placed.

The original data for the five lines are condensed in table 2. The number of bundle pairs (the primary double bundles) is given in parenthesis, and the number of intercalary bundles, if such are present, follows the + sign outside the parenthesis.

There are three outstanding features in this table.

First, the wide range of variation in the number and in the combinations of primary double bundles and intercalary bundles in both normal and abnormal plants observed when reasonably large series of seedlings are sectioned. It is clear that an anatomist who deals with only a few seedlings may obtain an altogether inadequate picture of the conditions which actually prevail in the species under investigation.

Second, notwithstanding the wide range of variation there are conspicuous modal classes in both normal and abnormal seedlings. In the normal plants these fall in all cases on four primary double bundles, without intercalary bundles, or with but one intercalary bundle; and in the trimerous plants, on six primary double bundles without intercalary bundles.

Third, the plants which are externally dimerous and trimerous are also clearly differentiated in internal morphology. The internal characters are, however, transgressive. It is impossible in some cases to distinguish from sections of the hypocotyl base alone between plants which superficially fall into the strictly alternative classes of dimery and trimery.

For purposes of more detailed analysis these formulae must be split up into their component elements.

A. Primary Double Bundles. The distribution of the number of primary double bundles in the five lines considered is shown in table 3 for dimerous and trimerous seedlings. These frequencies, reduced to a percentage basis, are represented graphically in figure 13. This shows that in all five lines the modal number of primary double bundles is two higher in the trimerous than in the dimerous plants. In the dimerous plants the modal class is in all cases 4; in the trimerous seedlings the modal class is 6. There is, therefore, a profound reorganization in the vascular anatomy of the seedling upon the assumption of a trimerous external organization.

Limiting our attention to primary double bundles and judging from modal classes only, an increase of fifty percent in the number of vascular elements is

TABLE 3. *Number of primary double bundles at base of hypocotyl in trimerous and dimerous seedlings*

	4	5	6	7	8	Total
Line 75						
Trimerous	1	11	121	8	1	142
Percent	0.70	7.75	85.21	5.63	0.70	
Dimerous	117	19	6	—	—	142
Percent	82.39	13.38	4.23			
Line 93						
Trimerous	1	18	132	4	—	155
Percent	0.65	11.61	85.16	2.58		
Dimerous	90	50	15	—	—	155
Percent	58.06	32.26	9.68			
Line 98						
Trimerous	1	11	170	1	—	183
Percent	0.55	6.01	92.90	0.55		
Dimerous	165	16	1	—	1	183
Percent	90.16	8.74	0.55		0.55	
Line 139						
Trimerous	1	8	97	—	—	106
Percent	0.94	7.55	91.51			
Dimerous	147	3	—	—	—	150
Percent	98.00	2.00				
Line 143						
Trimerous	5	46	159	9	2	221
Percent	2.26	20.81	71.94	4.07	0.90	
Dimerous	209	10	1	1	—	221
Percent	94.57	4.52	0.45	0.45		

FIG. 13. Percentage frequency distribution for number of primary double bundles at base of hypocotyl in dimerous (solid dots) and trimerous (circles) seedlings.

associated with an increase of fifty percent in the number of cotyledons and leaves. The distributions show, however, that this is only an incomplete and to some extent an erroneous, statement of the condition. In the dimerous seedlings the modal number of primary double bundles is 4, and all departures from the modal number are higher. In the trimerous seedlings the modal number is 6, and the departures may be in either the positive or the negative direction. The frequency distribution for the dimerous plants is therefore wholly skew, forming a typical J-curve; that for the trimerous plants more or less symmetrical,[8] but with departures occurring chiefly at smaller numbers of bundles.

The variation of primary double bundle number in dimerous and trimerous plants is, therefore, transgressive. The number of externally dimerous seedlings which might be considered to be anatomically trimerous, and the number of trimerous seedlings which might on anatomical grounds be considered dimerous is, however, very small.

Turning to the physical constants in table 4, we note that the mean

TABLE 4. *Statistical constants for number of primary double bundles at base of hypocotyl of trimerous plants and their normal controls*

	Mean	Standard Deviation	Coefficient of Variation
Line 75			
Trimerous (N = 142)........	5.98±.02	0.436±.017	7.28±.29
Dimerous (N = 142).........	4.22±.03	0.505±.020	11.97±.49
Actual difference...............	+1.76±.04	−0.069±.026	−4.69±.56
Relative difference..............	41.71	13.66	
Line 93			
Trimerous (N = 155)........	5.90±.02	0.396±.015	6.72±.26
Dimerous (N = 155).........	4.52±.04	0.666±.026	14.74±.58
Actual difference...............	+1.38±.04	−0.270±.030	−8.02±.63
Relative difference..............	30.53	40.54	
Line 98			
Trimerous (N = 183)........	5.93±.01	0.288±.010	4.86±.17
Dimerous (N = 183).........	4.12±.02	0.427±.015	10.36±.37
Actual difference...............	+1.81±.02	−0.139±.018	−5.50±.41
Relative difference..............	43.93	32.55	
Line 139			
Trimerous (N = 106)........	5.91±.02	0.323±.015	5.47±.25
Dimerous (N = 150).........	4.02±.01	0.140±.005	3.48±.14
Actual difference...............	+1.89±.02	+0.183±.016	+1.99±.28
Relative difference..............	47.01	130.71	
Line 143			
Trimerous (N = 221)........	5.81±.03	0.581±.019	10.01±.32
Dimerous (N = 221).........	4.07±.01	0.315±.010	7.75±.25
Actual difference...............	+1.74±.03	+0.266±.021	+2.26±.41
Relative difference..............	42.75	84.44	

[8] Line 139 is probably only an apparent exception to this rule. In both dimerous and trimerous seedlings variations from the modal class are extremely rare, and variations above the modal class have not been found in the 106 trimerous seedlings of this line sectioned.

umber of primary double bundles at the base of the hypocotyl of trimerous
lants is from 1.38 to 1.89 higher than in the dimerous controls.   This
epresents an excess of from 30.5 to 47.0 percent.

The five lines are not, however, consistent in the relative variability of
he normal and abnormal seedlings.

The standard deviation of the number of primary double bundles in the
rimerous plants is lower than that in the dimerous plants in lines 75, 93,
nd 98.   The differences are from 13.7 to 40.5 percent of the control values.
Lines 139 and 143 are in contrast to the foregoing.   The trimerous plants
f line 139 have a standard deviation of 0.323 ± .015 bundles, whereas the
limerous controls have a standard deviation of 0.140 ± .005, giving a
lifference of +.183 ± .016, which is 11.4 times as large as its probable error.
n line 143 the trimerous plants have a standard deviation of 0.581 ± .019
bundles as compared with 0.315 ± .010 bundles in the normal controls,
giving a difference of +.266 ± .021, which is 12.7 times as large as its
brobable error.   These are relative differences of +130.7 percent for line
39 and + 84.4 percent for line 143.

The same differences in variability between the lines is also conspic-
ious in the relative variabilities as measured by the coefficients of varia-
tion.   In the first three lines (75, 93, and 98) the coefficients of varia-
tion in the trimerous plants range from 4.9 to 7.3 percent as compared with
10.4 to 14.7 percent in the dimerous controls, giving differences in relative

TABLE 5.   *Number of intercalary bundles at base of hypocotyl in trimerous and dimerous seedlings*

	0	1	2	3	4	5	6	Total
Line 75								
Trimerous	116	20	5	—	—	—	1	142
Percent	81.69	14.08	3.52				0.70	
Dimerous	87	35	11	5	2	2	—	142
Percent	61.27	24.65	7.75	3.52	1.41	1.41		
Line 93								
Trimerous	129	21	4	—	1	—	—	155
Percent	83.23	13.55	2.58		0.65			
Dimerous	66	58	24	6	1	—	—	155
Percent	42.58	37.42	15.48	3.87	0.65			
Line 98								
Trimerous	165	16	1	—	—	1	—	183
Percent	90.16	8.74	0.55			0.55		
Dimerous	104	52	24	3	—	—	—	183
Percent	56.83	28.42	13.11	1.64				
Line 139								
Trimerous	96	10	—	—	—	—	—	106
Percent	90.57	9.43						
Dimerous	139	11	—	—	—	—	—	150
Percent	92.67	7.33						
Line 143								
Trimerous	157	64	—	—	—	—	—	221
Percent	71.04	28.95						
Dimerous	156	61	4	—	—	—	—	221
Percent	70.58	27.60	1.80					

variability ranging from −4.7 to −8.0 percent. In line 139 the coefficient of variation for trimerous seedlings is 5.47, whereas that for dimerous seedlings is 3.48. In line 143, the coefficient of variation for trimerous seedlings is 10.01, whereas that for dimerous seedlings is 7.75. Thus the relative variability in these two lines is greater in the *trimerous* than in the *dimerous* seedlings.

B. *Intercalary Bundles.* The distribution of the number of intercalary bundles (considered alone) in the base of the hypocotyl is shown in table 5.

FIG. 14. Percentage frequency distribution of number of intercalary bundles at base of hypocotyl in dimerous (solid dots) and trimerous (circles) seedlings.

The graphs in figure 14 show that for both dimerous and trimerous seedlings *no* intercalary bundles is the modal condition. In both cases the distribution is wholly skew. The normal and the abnormal seedlings of lines 75, 93, and 98 differ conspicuously, however, in that the percentage of seedlings with no intercalary bundles is much higher in the trimerous seedlings, while, conversely, the percentage of seedlings with from 1 to 5 intercalary bundles is much higher in the dimerous plants. These differences are not found in lines 139 and 143. As a matter of fact, the percentage of seedlings with no intercalary bundles is slightly, but perhaps not significantly, higher in the dimerous seedlings of line 139. In both lines 139 and 143 the number of seedlings with 1 or 2 intercalary bundles is very small indeed in both trimerous and dimerous series. The two lines are essentially alike in this regard and line 143 only is represented on the diagram.

The percentages of the seedlings with no intercalary bundles in the two lasses of plants and the differences in the percentage are as follows:

	Trimerous	Dimerous	Difference
Line 75	81.69	61.27	+20.42
Line 93	83.23	42.58	+40.65
Line 98	90.16	56.83	+33.33
Line 139	90.57	92.67	− 2.10
Line 143	71.04	70.58	+ 0.46

The physical constants in table 6 show that the mean number of inter-:alary bundles in both normal and abnormal seedlings is small—less than a ;ingle bundle per plant in every case.

TABLE 6. *Statistical constants for number of intercalary bundles at base of hypocotyl of trimerous plants and their normal controls*

	Mean	Standard Deviation	Coefficient of Variation
Line 75			
Trimerous (N = 142)	.25±.04	0.686±.027	270.69±42.86
Dimerous (N = 142)	.63±.06	1.024±.041	161.60±16.13
Actual difference	−.38±.07	−0.338±.049	+109.09±45.79
Relative difference	60.32	33.00	
Line 93			
Trimerous (N = 155)	.21±.03	.545±.021	255.80±36.78
Dimerous (N = 155)	.83±.05	.874±.033	105.79± 7.29
Actual difference	−.62±.06	− .329±.039	+150.01±37.50
Relative difference	74.69	37.64	
Line 98			
Trimerous (N = 183)	.13±.02	.480±.017	381.67±73.88
Dimerous (N = 183)	.60±.04	.776±.027	130.21± 9.62
Actual difference	−.47±.04	− .296±.032	+251.46±74.50
Relative difference	78.33	38.14	
Line 139			
Trimerous (N = 106)	.09±.02	.292±.014	309.84±64.50
Dimerous (N = 150)	.07±.01	.261±.010	355.48±70.95
Actual difference	+.02±.02	+ .031±.017	− 45.64±95.89
Relative difference	28.57	11.88	
Line 143			
Trimerous (N = 221)	.29±.02	.454±.015	156.62±12.21
Dimerous (N = 221)	.31±.02	.501±.016	160.44±12.76
Actual difference	−.02±.03	− .047±.021	− 3.82±17.66
Relative difference	6.45	9.38	

Again the lines fall into two classes, those in which the number of inter-calary bundles is conspicuously higher in the dimerous plants (lines 75, 93, and 98) and those in which the numbers are essentially identical (lines 139 and 143). In the trimerous seedlings of the first group the average number ranges from 0.13 to 0.25, whereas in the dimerous it varies from 0.60 to 0.83 bundle. Thus the mean number of intercalary bundles is

from 60 to 78 percent smaller in the trimerous than in the dimerous seedlings.

In line 139 the mean number of intercalary bundles is actually larger in the trimerous seedlings, but the difference is only $+.02 \pm .02$.

In line 143 the mean number of intercalary bundles in trimerous and dimerous seedlings is practically identical, the difference being only $-.02 \pm .03$. In both of these lines the differences are insignificant in comparison with their probable errors.

It is also interesting to note that in lines 75, 93, and 98 the differentiation between abnormal and normal seedlings is greater with respect to the number of intercalary bundles than with respect to primary double bundles. Turning back to table 4, we note that the number of primary double bundles is from 31 to 44 percent higher in the trimerous plants, whereas the number of intercalary bundles is from 60 to 78 percent lower. In lines 139 and 143 the difference in the mean of the number of primary double bundles of trimerous and dimerous plants is practically the same as in the other lines, but in these lines the two types of seedlings are essentially identical in number of intercalary bundles.

If we consider the comparative variability of dimerous and trimerous seedlings as to intercalary bundle number, we find that here, as in the case of number of primary double bundles, the lines differ among themselves. In all lines except 139 the standard deviations of number of intercalary bundles in the trimerous seedlings are smaller than in the dimerous. In lines 75, 93, and 98 the constants for the trimerous seedlings are from 33 to 38 percent smaller than those of the dimerous controls. In line 143 the difference has the same sign but is only -9.38 percent of the control value. In line 139 the difference is $+ 11.88$ percent.

The coefficients of variation are very high in both normal and abnormal seedlings, and this great variation renders the probable errors of little value as criteria of statistical significance of differences between the two types of seedlings. In lines 75, 93, and 98, the coefficients of variation for trimerous plants are conspicuously higher than those for the dimerous controls. In line 143 the coefficients of variation for the two types of seedlings are practically the same. In line 139, however, the coefficient of variation for the number of intercalary bundles is higher in dimerous than in trimerous plants.

C. Total Bundles. Having considered the frequency distribution and statistical constants for the two types of vascular structures found in the base of the hypocotyl, it is now desirable to combine the two types of bundles in order to consider the total number of vascular elements at this level.

This problem presents certain morphological difficulties. The primary double bundles are each derived from a single root pole, and do not become clearly divided into two bundles until the level of transition is reached from root structure to stem structure at the base of the hypocotyl. Many of the intercalary bundles appear at this level or a little lower. In determining

TABLE 7. *Total number of bundles at base of hypocotyl in trimerous and dimerous seedlings. Primary double bundles are counted as one bundle only*

	4	5	6	7	8	9	10	Total
ine 75								
Trimerous	—	I	115	21	3	I	I	142
Percent		0.70	80.99	14.79	2.11	0.70	0.70	
Dimerous	69	43	19	6	3	2	—	142
Percent	48.59	30.28	13.38	4.23	2.11	1.41		
ne 93								
Trimerous	—	5	130	18	2	—	—	155
Percent		3.23	83.87	11.61	1.29			
Dimerous	34	59	41	17	4	—	—	155
Percent	21.94	38.06	26.45	10.97	2.58			
ne 98								
Trimerous	—	4	166	12		I	—	183
Percent		2.19	90.71	6.56		0.55		
Dimerous	97	49	32	3	I	I	—	183
Percent	53.01	26.78	17.49	1.64	0.55	0.55		
ine 139								
Trimerous	—	5	96	5	—	—	—	106
Percent		4.72	90.57	4.72				
Dimerous	138	10	2	—	—	—	—	150
Percent	92.00	6.67	1.33					
ine 143								
Trimerous	2	18	165	30	5	I	—	221
Percent	0.90	8.14	74.66	13.57	2.26	0.45		
Dimerous	150	60	9	2	—	—	—	221
Percent	67.87	27.15	40	0.91				

FIG. 15. Percentage frequency distribution of total bundles at base of hypocotyl. Primary double bundles counted as single bundles.

the total number of bundles at the base of the hypocotyl, it therefore becomes a question as to whether we should count each primary double bundle as a single strand or as a double strand; adding, of course, the number of intercalary bundles in each case.

The distribution of total bundle number at this level according to the former method (primary double bundles counted as one, plus intercalaries) is shown in table 7, for both dimerous and trimerous seedlings. The results are shown clearly in figure 15.[9] The modal number is on 4 (lines 75, 98, 139, and 143) or 5 (line 93) bundles in the case of the dimerous seedlings, but invariably on 6 in the trimerous plantlets of the five lines. The distribution of number of bundles is almost wholly skew in the case of the normal seedlings, line 93 being slightly different from the others, but fairly symmetrical in the trimerous series.

The constants given in table 8 show that on the average the trimerous plants have from 0.77 to 1.91 bundles more than the dimerous plants. This is an excess of from 14.4 to 46.7 percent instead of the 50 percent which one might expect if the increase in number of bundles were proportional to the number of cotyledons or primordial leaves.

TABLE 8. *Statistical constants for total number of bundles at base of hypocotyl of trimerous plants and their normal controls. Primary double bundles are counted as one bundle only*

	Mean	Standard Deviation	Coefficient of Variation
Line 75			
Trimerous (N = 142)...........	6.23±.03	0.601±.024	9.65± .39
Dimerous (N = 142)...........	4.85±.06	1.087±.044	22.41± .94
Actual difference.................	+1.38±.07	−0.486±.050	−12.76±1.01
Relative difference.................	28.45	44.71	
Line 93			
Trimerous (N = 155)...........	6.11±.02	0.434±.017	7.10± .27
Dimerous (N = 155)...........	5.34±.06	1.019±.039	19.07± .76
Actual difference.................	+0.77±.06	−0.585±.042	−11.97± .80
Relative difference.................	14.41	57.41	
Line 98			
Trimerous (N = 183)...........	6.06±.02	0.365±.013	6.02± .21
Dimerous (N = 183)...........	4.72±.05	0.909±.032	19.28± .70
Actual difference.................	+1.34±.05	−0.544±.035	−13.26± .73
Relative difference.................	28.39	59.85	
Line 139			
Trimerous (N = 106)...........	6.00±.02	0.307±.014	5.12± .24
Dimerous (N = 150)...........	4.09±.02	0.334±.013	8.15± .32
Actual difference.................	+1.91±.03	−0.027±.019	− 3.03± .40
Relative difference.................	46.70	8.08	
Line 143			
Trimerous (N = 221)...........	6.10±.03	0.613±.020	10.06± .33
Dimerous (N = 221)...........	4.38±.03	0.609±.020	13.91± .45
Actual difference.................	+1.72±.04	+0.004±.028	− 3.85± .56
Relative difference.................	39.27	0.66	

[9] Lines 139 and 143 are in essential agreement with 75, 93, and 98, and are not drawn.

The variability, both absolute and relative, of the number of bundles is higher in dimerous than in trimerous plants. It is conspicuously higher in lines 75, 93, and 98. Thus the standard deviations for the trimerous plants range from 0.37 to 0.60 in the three lines as compared with 0.91 to 1.09 in the dimerous controls. The relative differences show that the variability of the trimerous plants is from 45 to 60 percent less than that of the dimerous plants. In the case of line 143, however, the difference between the standard deviation of the two types of seedlings is very small— less, indeed, than the probable error of the difference. Practically the same condition is found in line 139.

The coefficients of variation show that the trimerous plants have a variability in bundle number which is from 5.1 to 10.1 percent of the mean number of bundles, whereas the dimerous controls have a variability which is from 8.2 to 22.4 percent of the average number. In lines 75, 93, and 98 the difference between the two types is much more conspicuous than in lines 139 and 143.

Since in practically all cases, however, the primary double bundles have already clearly become two strands at the point where the intercalaries appear, it probably gives us a better conception of total bundle number here to count each primary bundle as *two*, and to add thereto the number of intercalaries. The actual and the percentage distribution according to this method are shown in table 9. Lines 75, 93, and 139 are represented in

TABLE 9. *Total number of bundles at base of hypocotyl in trimerous and dimerous seedlings. Primary double bundles are counted as two*

	8	9	10	11	12	13	14	15	16	17	Total
Line 75											
Trimerous..	—	—	2	8	109	12	10	—	1	—	142
Percent ..			1.41	5.63	76.76	8.45	7.04		0.70		142
Dimerous...	69	30	23	8	8	4	—	—	—	—	142
Percent ..	48.59	21.13	16.20	5.63	5.63	2.82					
Line 93											
Trimerous..	—	—	5	10	124	11	5	—	—	—	155
Percent ..			3.23	6.45	80.00	7.10	3.23				
Dimerous...	34	37	35	23	20	4	2	—	—	—	155
Percent ..	21.93	23.87	22.58	14.84	12.90	2.58	1.29				
Line 98											
Trimerous..	—	1	4	6	161	10	1	—	—	—	183
Percent ..		0.55	2.19	3.28	87.98	5.46	0.55				
Dimerous...	97	43	29	10	2	1	—	—	1		183
Percent ..	53.01	23.50	15.85	5.46	1.09	0.55			0.55		
Line 139											
Trimerous..	—	1	4	4	92	5	—	—	—	—	106
Percent ..		0.94	3.77	3.77	86.79	4.72					
Dimerous...	138	9	1	2	—	—	—	—	—	—	150
Percent ..	92.00	6.00	0.67	1.33							
Line 143											
Trimerous..	2	3	15	31	134	25	5	4	1	1	221
Percent ..	0.90	1.36	6.79	14.03	60.63	11.31	2.26	1.81	0.45	0.45	
Dimerous...	150	55	9	5	—	1	1	—	—	—	221
Percent ..	67.87	24.89	4.07	2.26		0.45	0.45				

TABLE 10. *Statistical constants for total number of bundles at base of hypocotyl of trimerous plants and their normal controls. Primary double bundles are counted as two*

	Mean	Standard Deviation	Coefficient of Variation
Line 75			
Trimerous (N = 142)	12.17±.04	0.750±.030	6.17±.25
Dimerous (N = 142)	9.07±.08	1.351±.054	14.90±.61
Actual difference	+3.10±.08	−0.601±.062	−8.73±.66
Relative difference	34.18	44.49	
Line 93			
Trimerous (N = 155)	12.01±.03	0.627±.024	5.22±.20
Dimerous (N = 155)	9.86±.08	1.483±.057	15.04±.59
Actual difference	+2.15±.08	−0.856±.062	−9.82±.62
Relative difference	21.81	57.72	
Line 98			
Trimerous (N = 183)	11.97±.02	0.495±.018	4.14±.15
Dimerous (N = 183)	8.84±.06	1.190±.042	13.47±.48
Actual difference	+3.13±.06	−0.695±.046	−9.33±.50
Relative difference	35.41	58.40	
Line 139			
Trimerous (N = 106)	11.91±.04	0.558±.026	4.69±.22
Dimerous (N = 150)	8.11±.02	0.440±.017	5.43±.21
Actual difference	+3.80±.04	+0.118±.031	−0.74±.30
Relative difference	46.85	26.82	
Line 143			
Trimerous (N = 221)	11.90±.05	1.105±.035	9.28±.30
Dimerous (N = 221)	8.45±.04	.831±.027	9.84±.32
Actual difference	+3.45±.06	+0.274±.044	−0.56±.44
Relative difference	40.83	32.97	

FIG. 16. Percentage frequency distribution of total bundles at base of hypocotyl in dimerous and trimerous seedlings. Primary double bundles counted as two.

figure 16.   Comparison of these figures with those in figure 15 shows essentially the same type of distribution for the dimerous and trimerous plants. The grades of the classes are merely about double what they were in the former method of treatment.

The statistical constants are compared in table 10.

FIG. 17.   Percentage frequency distribution of total bundles in central region of hypocotyl.

For all five lines the constants show a higher mean number of bundles in the trimerous than in the dimerous seedlings, the mean being approximately 12 in the former and 8 or 10 in the latter.   Thus the trimerous seedlings have from 21.8 to 46.9 percent more bundles than the dimerous seedlings.

Table 11. *Number of bundles in central region of hypocotyl of trimerous and dimerous seedlings*

	8	9	10	11	12	13	14	15	16	17	18	19	20	Total
Line 75														
Trimerous	1	3	5	36	292	40	29	5	1	4	—	—	—	416
Percent	0.24	0.72	1.20	8.65	70.19	9.62	6.97	1.20	0.24	0.96	—	—	—	
Dimerous	143	103	86	38	26	7	9	—	—	3	1	—	—	416
Percent	34.37	24.76	20.67	9.13	6.25	1.68	2.16	—	—	0.72	0.24	—	—	
Line 93														
Trimerous	—	—	8	32	382	82	38	12	1	—	1	—	1	557
Percent	—	—	1.44	5.75	68.58	14.72	6.82	2.15	0.18	—	0.18	—	0.18	
Dimerous	34	93	169	105	96	39	18	1	—	—	2	—	—	557
Percent	6.10	16.70	30.34	18.85	17.24	7.00	3.23	0.18	—	—	0.36	—	—	
Line 98														
Trimerous	—	1	6	12	297	21	8	—	—	—	—	—	—	345
Percent	—	0.29	1.74	3.48	86.09	6.09	2.32	—	—	—	—	—	—	
Dimerous	113	110	77	32	9	3	—	—	—	1	—	—	—	345
Percent	32.75	31.88	22.32	9.28	2.61	0.87	—	—	—	0.29	—	—	—	
Line 139														
Trimerous	—	—	4	8	84	6	3	1	—	—	—	—	—	106
Percent	—	—	3.77	7.55	79.25	5.66	2.83	0.94	—	—	—	—	—	
Dimerous	137	10	2	1	—	—	—	—	—	—	—	—	—	150
Percent	91.33	6.67	1.33	0.67	—	—	—	—	—	—	—	—	—	
Line 143														
Trimerous	2	1	11	14	136	21	25	6	3	1	1	—	—	221
Percent	0.90	0.45	4.98	6.33	61.54	9.50	11.31	2.71	1.36	0.45	0.45	—	—	
Dimerous	138	41	25	10	2	3	1	1	—	—	—	—	—	221
Percent	62.44	18.55	11.31	4.52	0.90	1.36	0.45	0.45	—	—	—	—	—	

In variability as measured by coefficient of variation, the dimerous plants exceed the trimerous throughout, conspicuously so in lines 75, 93, and 98. In their standard deviation, the dimerous also markedly exceed the trimerous in these three lines, but in lines 139 and 143 the trimerous plants slightly exceed the dimerous.

FIG. 18. Percentage frequency distribution of total bundles in central region of hypocotyl.

D. Summary for Base of Hypocotyl. For the base of the hypocotyl, therefore, it is evident that in total bundle number the trimerous seedlings decidedly exceed the dimerous ones. The intercalary bundles alone (which form but a small part of the total) tend to be more numerous in the dimerous seedlings.

In variability in bundle number at this region, dimerous seedlings in general exceed trimerous ones; although two of the five lines studied furnish slight exceptions to this rule.

FIG. 19. Percentage frequency distribution of total bundles in central region of hypocotyl.

3. *Central Region of Hypocotyl.* In the sections made in the central regions of the hypocotyl and of the epicotyl at both Cold Spring Harbor and Storrs, the total number of bundles was counted, no distinction being made between the bundles originating from the primary double bundles and those of intercalary origin.

FIG. 20. Percentage frequency distribution of total bundles in central region of hypocotyl.

The frequency distributions are shown in table 11. The relative frequencies for line 75 are shown in figure 17. The form of the distributions in line 98 is in essential agreement with those in line 75 and is not represented. The distributions for line 93 are represented in figure 18. The distribution for line 139 is shown in figure 19. That for line 143 appears in figure 20.

The conspicuous feature of these distributions is the wide variation in bundle number and the conspicuous skewness of the frequencies for the

normal plants of lines 75, 93, and 98. In these, bundle number ranges from 8 to 18 with a relatively large number of bundles in the lower classes. The modal number of bundles in the hypocotyl of normal seedlings of lines 75, 98, 139, and 143 is 8, while in line 93 it is 10.

The normal plants of the five lines differ conspicuously in variability. The number of seedlings falling in the modal class is relatively small and the range of variation relatively wide in lines 75, 93, and 98 as compared with line 139. Line 143 occupies an intermediate position in this regard.

In all the lines except 93 the distribution of number of bundles in the hypocotyl of normal seedlings is wholly skew, the frequency decreasing from the modal number (eight) towards the upper end of the range. In line 93 (figure 18) the distribution is also skew, but the frequency decreases from the modal number (ten) towards both ends of the range.

In the trimerous plantlets of all five series the modal number of bundles in the mid-region of the hypocotyl is 12. The extent of concentration into the modal class and the range of variation differs greatly in the five lines. This is very narrow in lines 98 and 139 and relatively wide in line 143.

The frequency distribution and figures bring out very clearly indeed the differentiation of the trimerous and dimerous seedlings in the number of vascular bundles.

TABLE 12. *Statistical constants for number of bundles in hypocotyl of trimerous and dimerous seedlings*

	Mean	Standard Deviation	Coefficient of Variation
Line 75			
Trimerous (N = 416)	12.19±.03	0.982±.023	8.06±.19
Dimerous (N = 416)	9.49±.05	1.645±.039	17.34±.42
Actual difference.	+2.70±.06	−0.663±.045	−9.28±.46
Relative difference	28.45	40.30	
Line 93			
Trimerous (N = 557)	12.29±.03	0.922±.019	7.50±.15
Dimerous (N = 557)	10.62±.04	1.525±.031	14.36±.30
Actual difference.	+1.67±.05	−0.603±.036	−6.86±.34
Relative difference	15.73	39.54	
Line 98			
Trimerous (N = 345)	12.03±.02	0.532±.014	4.42±.11
Dimerous (N = 345)	9.22±.04	1.197±.031	12.99±.34
Actual difference.	+2.81±.04	−0.665±.034	−8.57±.36
Relative difference	30.47	55.56	
Line 139			
Trimerous (N = 106)	11.99±.05	0.694±.032	5.78±.27
Dimerous (N = 150)	8.11±.02	0.409±.016	5.04±.20
Actual difference.	+3.88±.05	+0.285±.036	+0.74±.34
Relative difference	47.84	69.68	
Line 143			
Trimerous (N = 221)	12.29±.06	1.283±.041	10.44±.34
Dimerous (N = 221)	8.71±.05	1.187±.038	13.63±.45
Actual difference.	+3.58±.08	+0.096±.056	−3.19±.57
Relative difference.	41.10	8.09	

The differences between the lines can best be seen from the figures. For a more critical comparison we must have recourse to statistical constants and their probable errors.

The results for the hypocotyl of trimerous seedlings and their normal controls are set forth in table 12. Without exception the number of bundles in abnormal plants is higher than that in the control plants. The differences range from 1.7 to 3.9 bundles. These differences are many times as large as their probable errors and are unquestionably significant. The relative differences are about 16 percent in line 93, 30 percent in lines 75 and 98, 41 percent in line 143, and 48 percent in line 139.

Both the standard deviation and the coefficient of variation of the number of bundles in the hypocotyl are lower in the abnormal than in the normal plants in lines 75, 93, and 98. In lines 139 and 143 the relationship of the standard deviations of the trimerous and dimerous plants is exactly reversed, that of the trimerous plants being somewhat larger than that of the dimerous series. The difference in line 143 is $+.096 \pm .056$, which is nearly twice as large as its probable error and possibly statistically significant. In line 139 the difference in standard deviation is $+.285 \pm .036$. This difference is about 8 times as large as its probable error and unquestionably significant. The percentage differences in the standard deviations in lines 75, 93, and 98 range from -40 to -56 percent. In line 143 the percentage difference is $+8$ percent, while in line 139 it is $+70$ percent.

In line 143 the coefficient of variation is higher in dimerous plants (as it is in lines 75, 93, and 98), but in line 139 the trimerous show a slightly but perhaps not significantly higher relative variability.

The results as a whole show that the difference in the variability of bundle number in the two types of seedlings in lines 139 and 143 is not the same as that in lines 75, 93, and 98.

In interpreting these results we must remember that each primary double bundle at the base of the hypocotyl almost invariably divides to form two bundles at higher levels in the hypocotyl. Occasionally one of these branches may further divide into two. It is impossible in sections made in the central region of the hypocotyl to distinguish with certainty in every case between bundles originating through a division of the original protoxylem strands and those of intercalary origin.

The simplest working assumption is that the number of bundles in the central region of the hypocotyl will be given by twice the number of primary double bundles demonstrated at the base of the hypocotyl plus the number of intercalary bundles found at the base of the hypocotyl; or the number of bundles, b, at the central region should be given by

$$b = 2p + i$$

where p = primary double bundles and i = intercalary bundles.

A comparison of the number of bundles calculated by this formula with the number actually observed in the central region of the hypocotyl may be best made in a table of double entry. Table 13 gives the results for dimerous and table 14 the results for trimerous plants of line 93. The

TABLE 13. *Comparison of actual and theoretical number of bundles in hypocotyl of dimerous seedling*

Actual Number	8	9	10	11	12	13	14	Totals
Theoretical, $2p + i$ 8	12	13	6	3	—	—	—	34
9	—	14	17	3	1	1	1	37
10	—	1	22	6	5	1	—	35
11	—	—	1	9	9	3	1	23
12	—	—	—	1	14	4	1	20
13	—	—	—	—	—	1	3	4
14	—	—	—	—	—	—	2	2
Totals . .	12	28	46	22	29	10	8	155

TABLE 14. *Comparison of actual and theoretical number of bundles in hypocotyl of trimerous seedling*

Actual Number	10	11	12	13	14	15	20	Totals
Theoretical, $2p + i$ 10	1	1	3	—	—	—	—	5
11	—	6	3	—	1	—	—	10
12	—	2	102	12	6	1	1	124
13	—	—	—	8	3	—	—	11
14	—	—	—	—	4	1	—	5
Totals . .	1	9	108	20	14	2	1	155

frequencies for the cases in which the number of bundles at the mid-region of the hypocotyl calculated from the formula agrees with the number actually observed are printed in blackface type. The other lines give roughly comparable results.

It is clear that the number of hypocotyledonary bundles is not far from twice the number of primary root bundles plus the intercalary bundles. In rare cases the number of bundles in the hypocotyl is less than twice the root strands plus the number of intercalary bundles, since one of the root strands sometimes fails to divide. It may be, and not infrequently is, higher because of the appearance of extra intercalary bundles at a level higher than that sectioned at the base of the hypocotyl. In many cases the full complement of intercalary bundles has not appeared at this low level. In some cases it may be higher because of the secondary bifurcation above mentioned.

It is worth while to give the percentage frequencies of cases in which the number of bundles of the central region of the hypocotyl is given by the formula, and for comparison the relative number of cases in which it is in defect and in excess. The percentages are calculated from double entry tables like 13 and 14.

Trimerous Seedlings

	N	In Defect	$2p + i$	In Excess
ne 75	142	7.0	76.1	16.9
ne 93	155	1.3	78.1	20.7
ne 98	183	3.3	86.3	10.4
ne 139	106	7.6	80.2	12.3
ne 143	221	0.9	74.7	24.4

Dimerous Seedlings

	N	In Defect	$2p + i$	In Excess
ne 75	142	2.1	51.4	46.5
ne 93	155	1.9	47.7	50.3
ne 98	183	3.8	59.0	37.2
ne 139	150	0.7	98.7	0.7
ne 143	221	0.9	81.0	18.1

With the exception of the dimerous seedlings of line 139, the actually
observed number of bundles is in excess of the number given by the formula.

In lines 75, 93, and 98 the excess is far greater in dimerous than in tri-
merous seedlings. Thus in the dimerous class about 40 percent of the
seedlings show a number of bundles in the central region of the hypocotyl
which is in excess of twice the number of primary double bundles plus the
number of intercalary bundles at the base of the hypocotyl. In the case
of the trimerous seedlings the excess is much smaller, being roughly 20
percent. Thus it is clear that, *especially in the normal seedlings*, a large
number of the intercalary bundles do not extend to the base but appear in
the axis, ending blindly below, or that a considerable proportion of the
primary double bundles divide into more than two bundles.

In line 143 the number of cases in which the observed number of bundles
is greater than the calculated number is much more nearly equal in the two
types of seedlings. Thus in the trimerous seedlings 24.4 percent of the
seedlings have a number of bundles in the central region of the hypocotyl
greater than $2p + i$, whereas in the dimerous seedlings there are 18.1 per-
cent of seedlings of this class. In line 139 only 0.7 percent of the dimerous
seedlings show a number of bundles in excess of $2p + i$, whereas in the
trimerous seedlings 12.3 percent are in excess.

Thus lines 139 and 143 give results diametrically opposed to those of
the first three discussed.[10]

Summary for Central Region of Hypocotyl. It is evident from the
above statements that the number of bundles in the hypocotyl of trimerous
is decidedly higher than in that of dimerous seedlings; that in general the
bundle number is more variable in dimerous than in trimerous seedlings;
and that the intercalary bundles generally extend to a lower level in the
hypocotyl of trimerous than in that of dimerous seedlings.

[10] Note that the extremely small excess in line 139 may be due to the extraordinarily
normal character of the vascular system of the dimerous plants of this line.

TABLE 15. *Number of bundles in central region of epicotyl of trimerous and dimerous seedlings*

	10	11	12	13	14	15	16	17	18	19	20	21	22	Total
Line 75														
Trimerous	—	—	3	16	63	164	93	41	27	4	4	1	—	416
Percent			0.72	3.85	15.14	39.42	22.36	9.86	6.49	0.96	0.96	0.24		416
Dimerous	1	4	336	46	16	10	3	—	—					416
Percent	0.24	0.96	80.77	11.06	3.85	2.40	0.72							416
Line 93														
Trimerous			5	18	47	236	129	56	51	10	4		1	557
Percent			0.90	3.23	8.44	42.37	23.16	10.05	9.16	1.80	0.72		0.18	557
Dimerous	1	6	479	42	18	10	1	—	—					557
Percent	0.18	1.08	86.00	7.54	3.23	1.80	0.18							557
Line 98														
Trimerous			8	24	69	176	49	9	7	1	1	1	—	345
Percent			2.32	6.96	20.00	51.01	14.20	2.61	2.03	0.29	0.29	0.29		345
Dimerous			316	23	4	1	1	—	—	—		—	—	345
Percent			91.59	6.67	1.16	0.29	0.29							345
Line 139														
Trimerous				8	21	38	24	9	4	2				106
Percent				7.55	19.81	35.85	22.64	8.49	3.77	1.89				106
Dimerous			131	16	3	—	—	—	—	—				150
Percent			87.33	10.67	2.00									150
Line 143														
Trimerous			5	9	19	54	49	37	31	9	6	2	—	221
Percent			2.26	4.07	8.60	24.43	22.17	16.74	14.03	4.07	2.71	0.90		221
Dimerous			169	34	11	5	2	—	—	—		—	—	221
Percent			76.47	15.38	4.98	2.26	0.90							221

4. *Central Region of Epicotyl.* The frequency distributions of the umber of bundles occurring in the mid-region of the epicotyl appear in able 15 for both the abnormal and the control plants.   These distributions

FIG. 21.   Percentage frequency distribution of number of bundles in central region of epicotyl.

reduced to a percentage basis are represented graphically in figure 21 for line 75, in figure 22 for line 98, and in figure 23 for line 143.   The distributions

FIG. 22. Percentage frequency distribution of number of bundles in central region o
epicotyl.

for line 93 are essentially the same as those for line 75.   The graph for lin
139 is in essential agreement with that for line 98 and is not drawn.

In the dimerous plants the difference between the form of the frequency
distributions for number of epicotyledonary bundles in lines 75 and 93 o₁

ιe one hand and lines 98, 139, and 143 on the other is more apparent than
:al.   All five lines agree in showing the frequencies for the dimerous plants
rgely concentrated in a single modal class with a slight but evident skew-

FIG. 23.   Percentage frequency distribution of number of bundles in central region of
:picotyl.

ιess toward higher numbers of bundles.   In the case of lines 75 and 93
:here is a little over 1 percent of plants with fewer than the modal number
)f bundles, whereas in lines 98, 139, and 143 these do not occur in series of
:he numbers sectioned.   It is quite possible that the examination of a

larger series of plantlets would result in the finding of such seedlings in lines 98, 139, and 143, thus bringing the five series into full agreement.

In the trimerous seedlings the number of bundles shows rather wide, and fairly symmetrical, distribution about the modal class, which is 15 bundles. The lines differ, however, to a considerable extent in the amount of variation from the modal class. In lines 75, 93, and 98 the frequencies are to a far greater extent concentrated into the modal class, which contains from 39 to 51 percent of the frequencies, than in line 143, which contains only 24 percent of the cases. Line 139 is intermediate between these two extremes.

For a more precise comparison we utilize the constants set forth in table 16.

TABLE 16. *Statistical constants for number of bundles in epicotyl of trimerous and dimerous seedlings*

	Mean	Standard Deviation	Coefficient of Variation
Line 75			
Trimerous (N = 416)........	15.47±.04	1.355±.032	8.76±.21
Dimerous (N = 416)........	12.27±.02	0.735±.017	5.99±.14
Actual difference...............	+3.20±.04	+0.620±.036	+2.77±.24
Relative difference..............	26.08	84.35	
Line 93			
Trimerous (N = 557)........	15.65±.04	1.372±.028	8.77±.18
Dimerous (N = 557)........	12.19±.02	0.615±.012	5.05±.10
Actual difference...............	+3.46±.04	+0.757±.030	+3.72±.20
Relative difference..............	28.38	123.09	
Line 98			
Trimerous (N = 345)........	14.89±.04	1.152±.030	7.74±.20
Dimerous (N = 345)........	12.11±.02	0.416±.011	3.44±.09
Actual difference...............	+2.78±.04	+0.736±.032	+4.30±.22
Relative difference..............	22.96	176.92	
Line 139			
Trimerous (N = 106)........	15.24±.08	1.285±.060	8.44±.39
Dimerous (N = 150)........	12.15±.02	0.406±.016	3.35±.13
Actual difference...............	+3.09±.08	+0.879±.062	+5.09±.41
Relative difference..............	25.43	216.50	
Line 143			
Trimerous (N = 221)........	16.10±.08	1.750±.056	10.87±.35
Dimerous (N = 221)........	12.36±.03	0.757±.024	6.13±.20
Actual difference...............	+3.74±.09	+0.993±.061	+4.74±.40
Relative difference..............	30.26	131.18	

These results show that without exception the average number of bundles in the epicotyl is higher in trimerous than in dimerous seedlings. The difference ranges from 2.8 to 3.7 bundles. The probable errors of the differences are so small that there can be no reasonable doubt of their significance. In relative terms, the number of bundles in the abnormal plant is from 23.0 to 30.3 percent higher than that in the normal plant.

The variability of bundle number, both absolute and relative, is far higher in the abnormal (trimerous) plants. The relative differences show that the trimerous plants are from 84 to 217 percent more variable than the dimerous in the number of bundles in the central region of the epicotyl.

We now have to consider the relative number of bundles in the hypocotyl and in the epicotyl of the same plant. The constants for the normal plants are shown in table 17 and for the trimerous seedlings in table 18.

TABLE 17. *Comparison of statistical constants for number of bundles in hypocotyl and epicotyl of same plant. Seedlings with two cotyledons and two primordial leaves*

	Mean	Standard Deviation	Coefficient of Variation
ine 75 (N = 416)			
Hypocotyl...............	9.49±.05	1.645±.039	17.34±.42
Epicotyl...............	12.27±.02	0.735±.017	5.99±.14
Actual difference...............	+2.78±.05	−0.910±.043	−11.35±.44
Relative difference...............	29.29	55.31	
ine 93 (N = 557)			
Hypocotyl...............	10.62±.04	1.525±.031	14.36±.30
Epicotyl...............	12.19±.02	0.615±.012	5.05±.10
Actual difference...............	+1.57±.04	−0.910±.033	− 9.31±.32
Relative difference...............	14.78	59.67	
ine 98 (N = 345)			
Hypocotyl...............	9.22±.04	1.197±.031	12.99±.34
Epicotyl...............	12.11±.02	0.416±.011	3.44±.09
Actual difference...............	+2.89±.04	−0.781±.033	− 9.55±.35
Relative difference...............	31.34	65.24	
ine 139 (N = 150)			
Hypocotyl...............	8.11±.02	0.409±.016	5.04±.20
Epicotyl...............	12.15±.02	0.406±.016	3.35±.13
Actual difference...............	+4.04±.03	−0.003±.023	− 1.69±.24
Relative difference...............	49.82	0.73	
ine 143 (N = 221)			
Hypocotyl...............	8.71±.05	1.187±.038	13.63±.45
Epicotyl...............	12.36±.03	0.757±.024	6.13±.20
Actual difference...............	+3.65±.06	−0.430±.045	− 7.50±.49
Relative difference...............	41.91	36.23	

Normal and abnormal plants have in common a larger number of bundles in the epicotyl. The differences between the means for the two organs are clearly significant in comparison with their probable errors. The percentage differences show that the epicotyl has from 15 to 50 percent more bundles than the hypocotyl.

In the dimerous seedlings the variabilities, both absolute and relative, as measured by the standard deviation and coefficient of variation, are consistent in indicating a higher variability of bundle number in the hypocotyl. The difference is, however, very slight in line 139.

The difference between the variability of the hypocotyl and that of the epicotyl in the normal seedling as measured in terms of the standard devia-

tion is from 0.8 to 0.9 bundle, or from 55 to 65 percent of the larger value in lines 75, 93, and 98. In line 143 the difference is only 0.4 bundle, or 36 percent. In line 139 there is practically no difference in the standard deviation of bundle number in the mid-region of the first two internodes of the seedling.

TABLE 18. *Comparison of statistical constants for number of bundles in hypocotyl and epicotyl of same plant. Seedlings with three cotyledons and three primordial leaves*

	Mean	Standard Deviation	Coefficient of Variation
Line 75 (N = 416)			
Hypocotyl..................	12.19±.03	0.982±.023	8.06±.19
Epicotyl....................	15.47±.04	1.355±.032	8.76±.21
Actual difference...............	+3.28±.05	+0.373±.040	+0.70±.28
Relative difference..............	26.90	37.98	
Line 93 (N = 557)			
Hypocotyl..................	12.29±.03	0.922±.019	7.50±.15
Epicotyl....................	15.65±.04	1.372±.028	8.77±.18
Actual difference...............	+3.36±.05	+0.450±.033	+1.27±.22
Relative difference..............	27.33	48.81	
Line 98 (N = 345)			
Hypocotyl..................	12.03±.02	0.532±.014	4.42±.11
Epicotyl....................	14.89±.04	1.152±.030	7.74±.20
Actual difference...............	+2.86±.04	+0.620±.033	+3.32±.22
Relative difference..............	23.77	116.54	
Line 139 (N = 106)			
Hypocotyl..................	11.99±.05	0.694±.032	5.78±.27
Epicotyl....................	15.24±.08	1.285±.060	8.44±.39
Actual difference...............	+3.25±.09	+0.591±.068	+2.66±.48
Relative difference..............	27.11	85.16	
Line 143 (N = 221)			
Hypocotyl..................	12.29±.06	1.283±.041	10.44±.34
Epicotyl....................	16.10±.08	1.750±.056	10.87±.35
Actual difference...............	+3.81±.10	+0.467±.069	+0.43±.49
Relative difference..............	31.00	36.40	

Basing the comparisons on the coefficient of variation, we note that the coefficients for the hypocotyl range from 13.0 to 17.3 percent, whereas those for the epicotyl range from 3.4 to 6.0 percent in lines 75, 93, and 98. Thus there is a difference of about 10 percent in the coefficient of variation of bundle number in the hypocotyl and epicotyl (of the normal seedlings) of these lines. In line 143 this difference is only −7.50 percent. In line 139 it is only −1.69 percent.

The statistical relationship is in full accord with the anatomical findings recorded above (p. 68) where it was shown that the intercalary bundles of the hypocotyl as they approach the cotyledonary node fuse with the (normally 8) bundles originating by the division of the (normally 4) protoxylem poles of the primary root and completely lose their individuality, exactly six bundles emerging from the complex irrespective of the number which

ive entered it from the hypocotyl.[11]   Immediately above the cotyledons
e six remaining bundles approach, closing the cotyledonary gaps and form-
g a ring, the six members of which almost immediately divide, giving rise
the modal number, 12, which persists throughout the length of the
icotyl.   It is apparently the disappearance of the intercalary bundles as a
spicuous feature of the topography which results in the lowered variabil-
 of bundle number in the epicotyl as compared with the hypocotyl.

If this conclusion be true, we should find the least difference in the
riability of number of bundles in the central regions of the first two inter-
des in the lines in which intercalary bundles are least conspicuous as a
ature of the vascular topography.   As a matter of fact, this condition is
rongly supported by the results for the five lines investigated.   Turning
ck to table 6, showing the constants for number of intercalary bundles,
 note that lines 75, 93, and 98 have on the average from 0.60 to 0.83
tercalary bundle per (normal) plant.   These are the lines showing a
ative difference of 55 to 65 percent in the standard deviations as com-
red with 36 percent in line 143 with an average of 0.31 intercalary bundle,
d of only 0.73 percent for line 139 which has an average of only 0.07
tercalary bundle per plant.   The differences in the coefficients of variation
 hypocotyl and epicotyl are from −9.3 to −11.4 percent in the three lines
th from 0.6 to 0.8 intercalary bundle per plant, −7.5 percent in line 143
th an average of 0.31 intercalary bundle, and only −1.7 percent in line
9 with an average of only 0.07 intercalary bundle.

In the trimerous seedlings the relationship between the variation of the
imber of bundles in the hypocotyl and in the epicotyl is *just the reverse* of
at found in the normal type.   Variability as measured by the standard
viation is significantly higher in the epicotyl of all lines studied.   The
me is true if the coefficient of variation be used as a measure of variability,
though the differences for lines 75 and 143 are not large.

The anatomical explanation of this fact seems to be found in the pecu-
irities of behavior at the cotyledonary node.   As pointed out above
. 70), the epicotyledonary ring is typically made up of nine strands
stead of the six characteristic of the normal plant.   There is, therefore,
 the modal case an increase of fifty percent in the number of bundles in
e epicotyledonary ring of the trimerous plant as compared with the
merous plant.   Many of these bundles, but not all, divide to form the
undle system characteristic of the main course of the epicotyl.   It is this
riability in the extent of division of the bundles of the epicotyledonary
ng which, in connection with the low variability of the hypocotyl due to
e formation of but few intercalary bundles (except in lines 139 and 143,
here the number is about the same in normal and abnormal seedlings),
rings about the great variability in the bundle number of the mid-region

---

[11] This statement is based on a more detailed anatomical study of a portion of the
edlings.

of the epicotyl as compared with the mid-region of the hypocotyl, in the trimerous plants.

This condition furnishes an excellent example of the importance of a knowledge of descriptive morphology as an aid in interpreting biometric constants.

COMPARISON OF BUNDLE NUMBER IN THE FIVE LINES STUDIED

From the genetic standpoint it seems a matter of considerable interest to determine whether the three nominally pure lines[12] are differentiated with respect to their vascular anatomy.

A comparison of the percentage frequency distributions and the figures of the foregoing discussion will convince the reader that certain of the lines may be differentiated either in mean number of bundles, or in variability of number of bundles, or in both average number and variability of bundle number.

Since we hope to return to this problem later with even more extensive data, it seems unnecessary to consider the differences in the distributions and constants in detail at this time.

The results of this brief and superficial comparison seem to indicate that while different lines may not differ greatly in respect to certain of their vascular characters they may be differentiated with respect to others.

SUMMARY

This paper presents the results of a comparative and biometric study of the gross vascular anatomy of the seedling of *Phaseolus vulgaris*.

Two morphological types are considered: the normal, or dimerous seedling with two cotyledons and two primordial leaves, and the trimerous seedling with three cotyledons and three primordial leaves.

In normal seedlings, the vascular system of the root is typically tetrarch (with four protoxylem poles), and gives rise in the base of the hypocotyl to eight bundles which continue to the cotyledonary node. From the vascular complex at this point two strands are given off to each cotyledon and six are left, each of which divides into two to produce the typical twelve bundled condition of the epicotyl.

The trimerous seedlings typically possess six root poles instead of four, twelve bundles in the hypocotyl instead of eight, and nine primary epicotyledonary bundles instead of six. The nine primary epicotyledonary bundles may not all divide, however, so that the number of bundles in the central region of the epicotyl is variable, ranging in general from fourteen to eighteen.

[12] While the material employed in this study traces its origin from individual plant, the possibility of hybridization in the field is not excluded. Thus any comparison which may be made in this place must be regarded as preliminary merely.

In both types of seedlings, but more frequently in the normal ones, additional or intercalary bundles appear in the hypocotyl, either *de novo* or as a result of division of the primary strands.

The following constants[13] (table 19) for bundle number (at the different levels studied) epitomize the differences which characterize the two types of seedlings.

TABLE 19

	Trimerous Seedlings			Dimerous Seedlings		
	Mean	S. D	C. V.	Mean	S. D.	C. V.
Root poles......................						
Minimum.....................	5.02	.654	13.02	4.01	.081	2.03
Maximum.....................	5.16	.739	14.47	4.13	.338	8.18
Mean........................	5.09	.707	13.87	4.05	.171	4.19
Primary double bundles						
Minimum.....................	5.81	.288	4.86	4.02	.140	3.48
Maximum.....................	5.98	.581	10.01	4.52	.666	14.74
Mean........................	5.91	.405	6.87	4.19	.411	9.66
Intercalary bundles						
Minimum.....................	.09	.292	156.62	.07	.261	105.79
Maximum.....................	.29	.686	381.67	.83	1.024	355.48
Mean........................	.19	.491	274.92	.49	.687	182.70
Mid-region of hypocotyl						
Minimum.....................	11.99	.532	4.42	8.11	.409	5.04
Maximum.....................	12.29	1.283	10.44	10.62	1.645	17.34
Mean........................	12.16	.883	7.24	9.23	1.193	12.67
Mid-region of epicotyl						
Minimum.....................	14.89	1.152	7.74	12.11	.406	3.35
Maximum.....................	16.10	1.750	10.87	12.36	.757	6.13
Mean........................	15.47	1.383	8.92	12.22	.586	4.79

The variability of root pole number is distinctly higher in trimerous than in dimerous seedlings, because of the fact that in *all* seedlings a four-poled condition is characteristic of the main root system and prevails even in the trimerous forms up to within a few millimeters of the base of the hypocotyl. Sections in the upper root region in such seedlings therefore show a considerable number of four- and five-bundled individuals.

The number of intercalary bundles is highly variable in both seedling types. The standard deviation is distinctly larger in the dimerous forms, but because of the generally lower average number of intercalary bundles in trimerous seedlings, the relative variabilities as measured by the coefficient of variation are higher in the trimerous type.

In the central region of the hypocotyl the variability of bundle number, both absolute and relative, is far higher in the dimerous seedlings, due in large part to the generally higher standard deviation of the number of intercalary bundles in the dimerous type.

In the central region of the epicotyl just the reverse is true, the variability of bundle number being higher in the trimerous than in the dimerous seedling. This is evidently due to the facts (*a*) that the intercalary bundles

[13] Data for number of root poles are available for only three of the five lines.

of the hypocotyl are quite lost in the cotyledonary nodal vascular complex and thus do not affect the variability of the dimerous plants; and (b) that the doubling of the primary epicotyledonary bundles which almost invariably occurs in the normal seedling may not always take place, at least not at as low a level as the central region of the epicotyl, in the abnormal type.

Conclusions

The results of the foregoing morphological and biometric analyses justify the emphasis at this point of certain general considerations.

1. External differentiation such as that which characterizes dimerous and trimerous seedlings of *Phaseolus vulgaris* is accompanied by profound differences in internal structure.

2. Anatomical characters are by no means constant. On the contrary, they are very variable even in series of individuals which are genetically highly homogeneous. Morphological investigations based on limited series of individuals may, therefore, result in inadequate conceptions.

3. Variation in anatomical structure is not constant for the plant as a whole, but may differ from region to region or from organ to organ. Thus in the regions of the seedling here under consideration, hypocotyl and epicotyl differ widely in the variability of bundle number. Furthermore, differences in variability from organ to organ or from region to region are not constant, but may be conditioned by other morphological features. To illustrate from the case in hand, the variability of bundle number of normal seedlings is higher in the hypocotyl than in the epicotyl. In seedlings with three cotyledons and three primordial leaves, just the reverse is true. These differences in biometric constants are readily understandable in the light of a knowledge of comparative morphology.

4. The results of this study emphasize the importance of the use of both biometric and comparative methods to supplement each other in any attack upon the problems of general morphology or of morphogenesis.

THE VASCULAR ANATOMY OF NORMAL AND VARIANT SEEDLINGS OF PHASEOLUS VULGARIS

BY

J. ARTHUR HARRIS and EDMUND W. SINNOTT

Station for Experimental Evolution, Carnegie Institution of Washington

THE VASCULAR ANATOMY OF NORMAL AND VARIANT SEEDLINGS OF PHASEOLUS VULGARIS

BY J. ARTHUR HARRIS AND EDMUND W. SINNOTT

STATION FOR EXPERIMENTAL EVOLUTION, CARNEGIE INSTITUTION OF WASHINGTON

The investigations here summarized comprise a comparative and biometric study of the gross vascular anatomy of normal and variant seedlings of *Phaseolus vulgaris*.

Three morphological types have been considered, (*a*) the normal or *dimerous* seedling with two cotyledons and two primordial leaves, (*b*) the *trimerous* seedling with three cotyledons and three primordial leaves, and (*c*) the *hemitrimerous* seedling in which there are three cotyledons and two primordial leaves.

In normal seedlings, the vascular system of the root is typically tetrarch (with four protoxylem poles), and gives rise in the base of the hypocotyl to four pairs of double bundles which soon form a circle of eight bundles which continue to the cotyledonary node. At this point there is a complex vascular anastomosis. From it two strands are given off to each cotyledon. The remainder of the vascular tissue is reorganized into six strands, each of which typically soon divides into two, the twelve bundles thus formed comprising the vascular system of the epicotyl.

The trimerous seedlings typically possess six root poles instead of four, twelve bundles in the hypocotyl instead of eight, and nine primary epicotyledonary bundles instead of six. The nine primary epicotyledonary bundles do not all divide, however, so that the number of bundles in the central region of the epicotyl is variable ranging in general from fourteen to eighteen.

In both classes of seedlings, but more frequently in the normal type, additional or intercalary bundles appear in the hypocotyl, either *de novo* or as a result of division of the primary strands.

Four main groups of problems as to the vascular topography of these seedling types have been considered biometrically: First, the number of bundles at different levels in the seedling; second, the variability in bundle number; third, the differentiation in internal structure of seedlings which are externally dimerous, trimerous and hemitrimerous; and fourth, the interrelationship of bundle number in different regions of the seedling.

The following table of constants[1] summarizes the facts for number and variability of vascular bundles in various regions of the seedling and in the three types of seedlings.[2]

The constants in this table, and the frequency distributions from which the constants were computed, lead to the following conclusions.

	DIMEROUS SEEDLINGS			TRIMEROUS SEEDLINGS			HEMITRIMEROUS SEEDLINGS		
	Mean	S. D.	C. V.	Mean	S. D.	C. V.	Mean	S. D.	C. V.
Root poles									
Minimum	4.01	0.081	2.03	5.02	0.654	13.02			
Maximum	4.13	0.338	8.18	5.16	0.729	14.12			
Mean	4.05	0.171	4.19	5.09	0.707	13.87			
Primary double bundles									
Minimum	4.02	0.140	3.48	5.81	0.288	4.86	5.21	0.608	10.59
Maximum	4.52	0.666	14.74	5.98	0.581	10.01	5.74	0.750	14.07
Mean	4.19	0.411	9.66	5.91	0.405	6.87	5.49	0.676	12.37
Intercalary bundles									
Minimum	0.07	0.261	105.79	0.09	0.292	156.62	0.28	0.449	115.47
Maximum	0.83	1.024	355.48	0.29	0.686	381.67	0.53	1.148	214.68
Mean	0.49	0.687	182.70	0.19	0.491	274.92	0.44	0.737	163.82
Mid-region of hypo-cotyl									
Minimum	8.11	0.409	5.04	11.99	0.532	4.42	11.36	1.169	9.94
Maximum	10.62	1.645	17.34	12.29	1.283	10.44	12.32	1.524	12.87
Mean	9.23	1.193	12.67	12.16	0.883	7.24	11.94	1.307	10.96
Mid-region of epi-cotyl									
Minimum	12.11	0.406	3.35	14.89	1.152	7.74	12.93	1.245	9.07
Maximum	12.36	0.757	6.13	16.10	1.750	10.87	14.84	1.778	12.53
Mean	12.22	0.586	4.79	15.47	1.383	8.92	13.83	1.560	11.29

The modal number of primary double bundles in the region of transition from root to stem structure at the base of the hypocotyl is four in the dimerous and six in the trimerous and hemitrimerous seedling. In the normal seedlings more than four bundles may occur, but in no case have fewer than this number been observed. In the trimerous seedling variation both above and below the mode is found, the numbers ranging from four to eight. On the average the number is from 1.38 to 1.89 bundles higher (or from 30.5 to 47.0% higher) in the trimerous than in the dimerous seedlings.

Intercalary bundles, which are rather uncommon in seedling anatomy in general, occur in from 11 to 46% of the normal seedlings, whereas they are found in only 9 to 29% of the trimerous and in 28 to 43% of hemitrimerous seedlings. The average number of intercalary bundles is also generally higher in the dimerous plantlets.

Considering the total bundle number at the base of the hypocotyl (primary bundles plus intercalary bundles) the trimerous and hemitrimerous seedlings have from 0.77 to 1.91 bundles, or from 14.4 to 46.7% more than the dimerous seedlings. The differentiation of the dimerous

and trimerous seedlings is conspicuously shown by the frequency distributions of two of the lines shown in diagram 1.

In passing upward from the base of the hypocotyl, each primary bundle pair normally divides into two so that in the central region of the hypocotyl the bundle number is normally twice the number of primary double bundles at the base, plus the intercalary bundles. In many cases the number is somewhat in excess of this, however, showing either that new (intercalary) bundles have appeared or that some of the bundles have become subdivided.

The modal number of bundles in the mid-region of the hypocotyl is eight or ten in dimerous plantlets; in trimerous and hemitrimerous plantlets

DIAGRAM 1

Percentage frequency distributions of total bundles (primary double bundles counted as two) at the base of the hypocotyl in dimerous and trimerous seedlings of two lines. Abscissae represent bundle numbers, ordinates represent percentage frequencies.

it is twelve. On the average the number is from 1.7 to 3.8 bundles higher (or from 15.7 to 47.9% higher) in the trimerous than in the dimerous seedlings. The differentiation of the two classes of seedlings in their vascular anatomy at the level is clearly shown in diagram 2.

The bundles in the mid-region of the epicotyl show in dimerous plantlets a modal number of twelve, whereas in trimerous seedlings it is fifteen. On the average there are from 2.8 to 3.7, or from 23.0 to 30.2%, more bundles in the epicotyl of the trimerous than in the dimerous seedling.

The form of the frequency distributions for two of the lines is shown in diagram 3. The epicotyl of the hemitrimerous is in essentials of anatomy identical with that of the dimerous seedling.

Not only are there marked differences in the actual number of bundles, but the variability of bundle number changes from region to region of the seedling, and differs in the three seedling types. Whether judged by range, standard deviation or coefficient of variation, the variability of bundle number in the central region of hypocotyl tends to be distinctly higher in the dimerous than in the trimerous plantlets; but in the epicotyl just the reverse is true, the variability of the trimerous plantlets exceeding that of the dimerous. These differences are conspicuous in diagrams 2

DIAGRAM 2

Percentage frequency distribution of number of bundles in central region of hypocotyl in dimerous and trimerous seedlings. Abscissae represent bundle numbers, ordinates represent percentage frequencies.

and 3. In the first case it is the dimerous plantlets, in the second case it is the trimerous ones which show the greater variability. Apparently this is due to differences in the number of intercalary bundles in the hypocotyl, and to the extent of division of the bundles in the epicotyl, of the two types of seedlings.

The coefficients of correlation between various bundle systems also differ widely. In both trimerous and dimerous seedlings there is a negative correlation between the number of primary double bundles and the number of intercalary bundles at the base of the hypocotyl. Thus the number of intercalary bundles is smaller in seedlings with larger numbers of primary double bundles and vice versa. This result for seedlings of the

same (external) morphological type is in agreement with those obtained by a comparison of seedlings which are externally dimerous and trimerous, since the latter frequently have a larger number of primary double bundles but a smaller number of intercalary bundles than the former. In both types of seedlings variation in the number of intercalary bundles is the primary factor in determining variation in the total number of bundles at the base of the hypocotyl.

Turning to the problem of the interrelationship of bundle number at different levels in the seedling we find that there is a substantial correla-

DIAGRAM 3

Percentage frequency distributions of total bundle number in the central region of the epicotyl of dimerous and trimerous seedlings of two lines. Abscissae represent bundle numbers, ordinates represent percentage frequencies.

tion between the numbers of the three classes of bundles—primary double bundles, intercalary bundles, and total bundles—at the base of the hypocotyl and the number of bundles in the central region of the hypocotyl. In the normal seedlings the coefficients average $+0.509$ for number of primary double bundles and number of hypocotyledonary bundles, $+0.629$ for intercalary bundles and hypocotyledonary bundles, and $+0.813$ for total bundles and hypocotyledonary bundles. In the trimerous plants these correlations average $+0.381$, $+0.238$ and $+0.598$, respectively. The correlations for normal plantlets are practically without exception higher than those for abnormal seedlings.

The correlations between the number of bundles in the hypocotyl (both basal region and central region) on the one hand and the number of

CORRELATION OF MORPHOLOGICAL VARIATIONS IN THE SEEDLING OF PHASEOLUS VULGARIS

J. ARTHUR HARRIS AND B. T. AVERY

NEW YORK
1918

Correlation of morphological variations in the seedling of Phaseolus vulgaris

J. Arthur Harris and B. T. Avery

Introductory remarks

During the past several years one of us has had under way extensive experiments on the differential death-rate of bean seedlings. Individuals differing in structure also differ in their capacity for survival under field conditions,* and in such physiological characteristics as capacity for the development of the tissues of the primordial† and of the subsequent leaves.‡

Some tens of thousands of seedlings of known morphological characteristics have been exposed to risk, as the life insurance statisticians express it, in an attempt to determine the selective value of the various morphological variations. These seedlings were, for technical reasons, necessarily planted in the field at a time when the cotyledonary node and the primordial node only could be studied. It is evident that the capacity of the plant for survival may be in some degree dependent upon characters developed later in ontogenesis, but correlated with characters of the first or second node of the seedling.

However this may be, it is certainly true that a full knowledge of the morphology and physiology of the variant bean seedling demands a thoroughgoing investigation of the correlation between the structure of the first two leaf whorls and that of later whorls. We have, therefore, been forced to consider the problem of the morphological character of the leaf whorls produced at the third

---

* Harris, J. Arthur. A simple demonstration of the action of natural selection. Science II. **36**: 713–715. 1912.

† Harris, J. Arthur. Studies on the correlation of morphological and physiological characters: the development of the primordial leaves in teratological bean seedlings. Genetics **1**: 185–196. 1916.

‡ Harris, J. Arthur. Further studies on the interrelationship of morphological and physiological characters in seedlings of *Phaseolus*. Mem. Brooklyn Bot. Gard. 1. *In press*.

node in the case of plants showing various structural abnormali[t]
at the first two nodes.

Phaseolus is well suited for such investigations. The nor[m]
seedling has two cotyledons, inserted at the same level, and t[
opposite primordial leaves. A large number of structural va[ria-]
tions, four types of which will be considered in this paper, m[
occur. The chief disadvantage lies in the rarity of many of [
variations in the lines with which we have dealt. The securing [
adequate series is excessively laborious. The present pape[r]
based upon a careful study of the variations in the first three no[de]s
of 16,348 plants, which were selected from about 450,000 seedli[
examined for the characters of the first and second node.

When in the following paragraphs we refer to normal a[
abnormal plants or seedlings, it must be understood that t[
applies to the characteristics of the individual as determined [
the basis of the first two nodes, the cotyledonary and the p[ri-]
mordial only. In its later development the "abnormal" pl[
may remain "abnormal" or become "normal," and the "norm[
plant may either continue to be "normal" or become "abnorm[

The nature and method of classification of the abnormali[
dealt with will be discussed in the presentation of the data belo[

MATERIALS AND METHODS

The materials upon which this study is based are a serie[
lines of White Navy beans grown at the Station for Experime[
Evolution during the past several years. The seeds were [har-]
vested from field cultures in 1915 and germinated in sand in [
autumn of 1916.

Seedlings which were abnormal in the characters of the [fir]t
or second node, i. e., in the number or insertion of the cotyled[s]
or of the primordial leaves, were sorted out for potting in soil [
subsequent study of the third node, that normally giving ris[
the first compound leaf.

For each abnormal individual, a normal control seedling f[
the same parent plant was taken at random to serve as a basi[
comparison. Both were potted in soil and grown to a stage w[
the characteristics of the third node could be accurately de[
mined.

or the onerous preliminary examination of nearly a half a million seedlings we are greatly indebted to Miss Edna K. Lockwood, Miss Margaret Gavin and especially to Miss Lillie Gavin.

PRESENTATION AND ANALYSIS OF DATA

The slightest abnormality which we have been able to discover occurring in considerable numbers of bean seedlings is the vertical separation of the two normally opposite cotyledons. So imperceptible is the line of transition between normal and abnormal that personal equation must play some part in classification. The cotyledons may be much more widely separated. The variation is a purely graduated one, with no sharp lines of demarcation between the different degrees of separation. Generally we have recognized three grades, but because of the rarity of plants with more widely separated cotyledons we have in this paper grouped our data into two classes only. The first comprises plants with cotyledons 2–3 mm. apart. The second includes all those in which they are more distant.

The number as well as the position of the cotyledons may vary. Plants with three instead of two cotyledons fall into two groups; those with the normal pair of primordial leaves and those with a whorl of three leaves. The latter are by far the more abundant.

Abnormality developed subsequently to the selection of the seedlings in the preliminary sorting may affect either the internode between the second and the third nodes, that is, between the primordial leaf whorl and the point of insertion of the first compound leaf or leaf whorl, or it may be confined to the number or structure (or number and structure) of the leaves inserted at the third node.

In the original selection of individuals abnormal in the characters of the first or second node, only those with sensibly normal axes (hypocotyl and epicotyl) were chosen for the purposes of the present study.

Two types of abnormality in the axis beyond the second (the primordial) node have been considered.

The first is a sensible broadening of the axis, identical with or similar to fasciation. This is a graduated character. The line of demarcation between normal and abnormal is not clearly marked.

and personal equation may influence in some degree the classifi
tion of the seedlings.

The second is a division of the axis into two coördinate branc
each with a terminal bud.

The frequencies of the two types of axial variation are too sm
to justify detailed discussion. The entries in TABLE I show t

TABLE I

FREQUENCIES OF ABNORMALITY OF SECOND INTERNODE IN NORMAL AND ABNOR
SEEDLINGS

Class of abnormality	Actual frequencies			Percentage frequencies		
	Normal inter-node	Broad-ened inter-node	Di-vided inter-node	Normal internode	Broadened internode	Divi inter:
Two cotyledons slightly separated; two primordial leaves..........	4,017	5	8	99.6774	.1240	
Normal control.................	4,029	0	1	99.9751	......	
Difference.....................	−12	+ 5	+ 7	−.2977	+ .1240	+.
Two cotyledons widely separated; two primordial leaves..........	878	2	1	99.6594	.2270	
Normal control.................	881	0	0	100.0000	......	
Difference.....................	−3	+ 2	+ 1	− .3406	+ .2270	+.
Three cotyledons; two primordial leaves	813	12	0	98.5454	1.4546	
Normal control.................	825	0	0	100.0000	......	
Difference.....................	−12	+12	0	−1.4546	+1.4546	
Three cotyledons; three primordial leaves......................	2,410	14	14	98.8515	.5742	
Normal control.................	2,436	2	0	99.9179	.0820	
Difference.....................	−26	+12	+14	−1.0664	+ .4922	+.

in every instance in which any individuals at all are available
seedlings which are abnormal in either the cotyledonary or
primordial node show a higher percentage of abnormality in
structure of the internode beyond the second node than do
normal controls.

We now turn to a consideration of variation in the leave
serted at the third node. The leaves of plants with abnorm
of the axis should not be combined with those having normal
They are not sufficiently numerous for separate consideration

Confining our attention to seedlings which have a normal
for at least the length of the second internode of the epicoty

ha\ the frequencies shown in TABLES II–V.* The character of the
con·ol plants is also given.

TABLE II

SEDLINGS WITH TWO COTYLEDONS SLIGHTLY SEPARATED AND TWO PRIMORDIAL
LEAVES

Nober of leav per node	Actual frequencies		Percentage frequencies		Difference
	Abnormal	Control	Abnormal	Control	
I	3,791	3,853	94.37	95.63	−1.26
2	225	176	5.60	4.37	+1.23
3	I	0	.02	.00	+ .02
Tots......	4,017	4,029	99.99	100.00	

TABLE III

SEDLINGS WITH TWO COTYLEDONS WIDELY SEPARATED AND TWO PRIMORDIAL
LEAVES

Nnber of leav per node	Actual frequencies		Percentage frequencies		Difference
	Abnormal	Control	Abnormal	Control	
I	811	840	92.37	95.35	−2.98
2	67	41	7.63	4.65	+2.98
Tots......	878	881	100.00	100.00	

TABLE IV

SEEDLINGS WITH THREE COTYLEDONS AND TWO PRIMORDIAL LEAVES

Nnbes of leav per node	Actual frequencies		Percentage frequencies		Difference
	Abnormal	Control	Abnormal	Control	
I	591	792	72.69	96.00	−23.31
2	221	33	27.18	4.00	+23.18
3	I	0	.12	.00	+ .12
Tots......	813	825	99.99	100.00	

TABLE V

SEEDLINGS WITH THREE COTYLEDONS AND THREE PRIMORDIAL LEAVES

Nmber of leav per node	Actual frequencies		Percentage frequencies		Difference
	Abnormal	Control	Abnormal	Control	
I	1,632	2,200	67.72	90.31	−22.59
2	771	236	31.99	9.69	+22.30
3	7	0	.29	.00	+ .29
Tots......	2,410	2,436	100.00	100.00	......

In these tables the numbers of control plants are not exactly identical with the
nubers of abnormal plants, since some of those selected as normal in the seedling
sta. showed abnormality of the axis in subsequent development and are omitted
her where we are discussing abnormalities of foliar characters only.

In each of the types of abnormality dealt with the abnor al
series show a higher proportion of the individuals with two or t se
leaves at the third node than do their normal controls.

Furthermore, seedlings showing different types of abnorma ty
at the first nodes also differ among themselves in the exter of
abnormality at the third node. Thus plants which are nor al
except for slight separation of the cotyledons have two or t ee
leaves at the second node instead of the single leaf normally fo id
in 5.63 per cent. of the individuals. Plants with the cotyle ns
more widely separated have 7.63 per cent. of their number with o
or three instead of a single leaf.

When one turns to the groups of plants which have thre n-
stead of two cotyledons, a conspicuous difference is at once appa it.
Plants which have three cotyledons and a normal pair of primo al
leaves produce two or three instead of a single leaf at the t rd
node in 27.31 per cent. of the cases. Seedlings with three c y-
ledons and a whorl of three primordial leaves instead of the no al
pair at the third node have 32.29 per cent. of the individuals th
two or three leaves at the third node.

Heretofore the number of leaves inserted at the third node as
furnished the only measure of variation at this region of the is.
We now propose to consider variation in the organization of he
leaves themselves. It will not be possible to do this in the d ail
in which we hope to treat the problem ultimately. The ran of
variation in the division of the bean leaf is rather great, anc he
laws governing it are doubtless very complicated. Some prog ss
has already been made on the problem, but for the present we all
limit our discussion to the number of leaflets, leaving the prol m
of their arrangement for treatment when even larger series of ta
are at our disposal.

The actual frequencies of number of leaflets per leaf prod ed
at the third node are shown in TABLE VI.

The most conspicuous feature of this table is the bim lal
nature of the distribution. The modes are on three and six, is
to be expected from the fact that the distribution of the w le
number of leaflets depends upon plants with from one to t ee
leaves at the third node.

Because of the wide range of variation in leaflet number is

ot feasible to reduce these frequencies for the individual classes
o a percentage basis for comparisons. This has, however, been
lone for larger groups secured by combining all the seedlings

TABLE VI

NUMBER OF LEAFLETS PRODUCED AT THE THIRD NODE BY SEEDLINGS OF VARIOUS
TYPES

Number of leaflets	Two cotyledons slightly separated and two primordial leaves		Two cotyledons widely separated and two primordial leaves		Three cotyledons and two primordial leaves		Three cotyledons and three primordial leaves	
	Abnormal	Control	Abnormal	Control	Abnormal	Control	Abnormal	Control
I	5	2	I	. . .	I	. . .	. . . .	. . . .
2	16	6	4	. . .	3	2	2	5
3	3,741	3,825	801	835	572	782	1,602	2,185
4	27	21	5	4	14	5	28	8
5	10	2	I	2	6	3	9	2
6	215	173	65	40	203	33	751	236
7	I	. . . .	I	. . .	12	. . .	12	. . . .
8	. . . .	. . . .	. . .	. . .	I	. . .	. . . .	. . . .
9	I	. . . .	. . .	. . .	I	. . .	3	. . . .
10	I	. . . .	. . .	. . .	. . .	. . .	2	. . . .
11	. . . .	. . . .	. . .	. . .	. . .	. . .	I	. . . .
Totals	4,017	4,029	878	881	813	825	2,410	2,436

howing merely separation of the cotyledons and all those showing
hree cotyledons instead of the normal two. The results are shown
n the accompanying TABLE VII, which is self explanatory.

TABLE VII

COMPARISON OF THE NUMBER OF LEAFLETS IN DICOTYLEDONOUS AND TRICOTYLEDO-
NOUS SEEDLINGS WITH THAT IN THEIR NORMAL CONTROLS

Number of leaflets	Seedlings with cotyledons separated		Seedlings with three cotyledons	
	Abnormal	Control	Abnormal	Control
I	.12	.04	.03	. . . .
2	.41	.12	.16	.21
3	92.79	94.91	67.45	90.98
4	.65	.51	1.30	.40
5	.22	.08	.47	.15
6	5.72	3.52	29.60	8.25
7	.04	.81	.74	. . . .
8	. . . .	. . . .	.03	. . . .
9	.02	. . . .	.12	. . . .
10	.02	. . . .	.06	. . . .
11	. . . .	. . . .	.03	. . . .

A comparison may be made without the combination of dif-
erent grades of abnormality by grouping the number of leaflets

around the modal classes 3, 6 and 9.   The results in TABLE VIII
show essentially the same relationships as those given in TABLES
VI–VII.   First, the higher leaflet numbers are more extensively
represented in the abnormal plants of each of the four types than
they are in the controls.   Second, the tricotyledonous plants show
a far greater increase in the number of leaflets inserted at the third
node than do those abnormal only in the position at which the
two cotyledons are inserted.

TABLE VIII

PERCENTAGE FREQUENCIES OF NUMBERS OF LEAFLETS IN SEEDLINGS OF VARIOUS
TYPES

Class of abnormality	Number of leaflets		
	1–4	5–7	8–11
Two cotyledons slightly separated and two primordial leaves...	94.32	5.62	.05
Control...	95.66	4.34	....
Difference...	− 1.34	+ 1.28	+.05
Two cotyledons widely separated and two primordial leaves...	92.37	7.63	....
Control...	95.23	4.77	....
Difference...	− 2.86	+ 2.86	....
Three cotyledons and two primordial leaves...	72.57	27.18	.24
Control...	95.64	4.36	....
Difference...	−23.07	+22.82	+.2.
Three cotyledons and three primordial leaves...	67.72	32.03	.2:
Control...	90.23	9.77	....
Difference...	−22.51	+22.26	+.2

In substantiation of these conclusions the reader will note th
in the class with slightly separated cotyledons 5.68 per cent. of t
plants have from five to ten leaflets as compared with 4.34 p
cent. with five and six leaflets in the control series.   In seedlin
with more widely separated cotyledons but no other abnormal
there are 7.63 per cent. of the plants with five to seven leaflets
compared with 4.77 per cent. of the normal controls with five
six leaflets.   Seedlings with three cotyledons but the normal numb
of primordial leaves have 27.43 per cent. of the individuals wi
from five to nine leaflets as compared with 4.36 per cent. with f
or six leaflets in the normal controls.

Plants with a trimerous cotyledonary and primordial whl
have 32.28 per cent. of the seedlings with from five to eleven leaf s

a compared with 9.77 per cent. with five and six leaflets in the normal controls.

Taking the average number of leaflets per plant as a basis of comparison between the abnormal plants and their controls we find the results in TABLE IX.

TABLE IX

MEAN NUMBER OF LEAFLETS IN SEEDLINGS OF VARIOUS TYPES

Class of abnormality	Mean number of leaflets in abnormals	Mean number of leaflets in controls	Difference
Two cotyledons slightly separated; two primordial leaves	3.170	3.133	+0.037
Two cotyledons widely separated; two primordial leaves	3.228	3.191	+0.037
Three cotyledons; two primordial leaves	3.847	3.131	+0.716
Three cotyledons; three primordial leaves	3.990	3.294	+0.606

Note (a) that for each type of normality the average number of leaflets is greater in the abnormal individuals than in the normal, and (b) that the difference between the abnormal class and its control is far greater in the case of the plants with three cotyledons than in those in which the abnormality in the cotyledonary whorl consists merely in the separation of the two cotyledons.

Thus the results for number of leaflets substantiates the conclusion based upon number of leaves.

Evidently, however, the number of leaflets is to a great extent determined by the number of leaves. The problem now arises: Are there differences in the average number of leaflets per leaf in the abnormal and normal individuals?

Means and their differences have been determined, but are so slight that conclusions must be deferred until further series of data are available.

Just one other method of dealing with the problem of the correlation in structural variation may now be considered.

Number of leaflets is, in the materials dealt with, practically an integral variate. In examining a large series of plants those with partial division of a leaflet, representing transition stages between a leaf with n and one with $n + 1$ leaflets are sometimes found. Such cases are, however, relatively rare. The lobing of the leaflet has therefore been disregarded in the foregoing treat-

ment.   A leaf with three leaflets, one of which has a lobe, has bee
recorded as 3 in the tables, not as an intermediate between thre
and four.   This has simplified the tabling of the data, and the ca
culation of the simple constants necessary to the interpretatio
the data, without any material loss in accuracy.

One may, however, inquire whether there are differences in th
degree of lobing of the leaflets produced at the third node in plant
which are normal and in plants which are abnormal in the cha
acters of the first and second node.   Because of the very low pe
centages of lobing in the leaflet no stress whatever is to be lai
upon the exact values found, even in samples containing severa
hundreds or thousands of plants, because of the great difficultie
of determining the probable error of a small percentage.

The results are given in Table X.

TABLE X

PERCENTAGE OF LOBING IN THE LEAVES OF SEEDLINGS OF VARIOUS TYPES

Class of abnormality	Abnormal plants	Control plants	Differer
Two cotyledons slightly separated; two primordial leaves....................................	0.349	0.273	+0.07
Two cotyledons widely separated; two primordial leaves....................................	0.456	0.112	+0.34
Three cotyledons; two primordial leaves	1.599	0.242	+1.35
Three cotyledons; three primordial leaves	0.954	0.328	+0.62

Here the percentage frequency of plants with one or two lob
on the leaflets* are given for each type of abnormality dealt wit
and compared with that found in the control series.

Two relationships seem clearly indicated by the constants
this table.

First, the tendency to the production of lobes is greater in t
leaflets produced by abnormal plants of all four types than
their normal controls.

Second, the tendency to the production of lobes is greater
the leaflets of plants with a trimerous cotyledonary whorl, an
either a dimerous or trimerous primordial whorl, than it is
plants in which the sole abnormality consists in the separation
the two cotyledons in their insertion on the axis of the plant.

* In the case of two lobes both may occur on the same leaflet or they may be
different leaflets.

RECAPITULATION

This paper presents the results of a first attempt to determine some of the correlations in the structural variations of the seedling of *Phaseolus vulgaris*.

The materials are drawn from a series of lines of Navy beans grown for the past several years at the Station for Experimental Evolution. The seeds used were harvested from plants of selected ancestry. Neither of these factors will, we believe, invalidate the conclusions drawn in this paper. These conclusions will not necessarily apply to certain entirely abnormal races.

Fasciation-like broadening of the axis and longitudinal division of the axis distal to the insertion of the primordial leaves are both more frequent in seedlings showing separation of the cotyledons and in tricotyledonous seedlings than in those which are normal.

Seedlings which are normal except for the separation of the cotyledons and those which have three cotyledons and a normal pair of primordial leaves or three cotyledons and a whorl of three primordial leaves produce a larger number of leaves, a larger number of leaflets and a higher percentage of leaves with lobes at the third node than do those which are normal in their cotyledonary and primordial leaf characters.

Seedlings which are tricotyledonous, with either a normal pair or a whorl of three primordial leaves, show higher percentages of variation in the axis, or the leaves produced by the axis, distal to the primordial leaves than do those which are normal except for the separation of the two cotyledons.

These studies will be continued.

STATION FOR EXPERIMENTAL EVOLUTION,
    COLD SPRING HARBOR, NEW YORK

FURTHER STUDIES ON THE INTERRELATIONSHIP OF MORPHOLOGICAL AND PHYSIOLOGICAL CHARACTERS IN SEEDLINGS OF PHASEOLUS

J. ARTHUR HARRIS

FURTHER STUDIES ON THE INTERRELATIONSHIP OF MORPHOLOGICAL AND PHYSIOLOGICAL CHARACTERS IN SEEDLINGS OF PHASEOLUS [1]

J. ARTHUR HARRIS

Station for Experimental Evolution, Cold Spring Harbor, N. Y.

Introductory Remarks

In a series of papers published during the past several years I have emphasized the importance of investigations of the relationship between the morphological and the physiological characteristics of the organ and of the organism.

The structural variations of the organs of which the organism is made up are the resultant of intrinsic and extrinsic factors—of heredity and environment, or of nature and nurture. Morphogenetic processes must, therefore, be investigated by physiological methods, and be interpreted in physiological, and ultimately in physical and chemical, terms.

The purpose of this paper is to supplement and extend the results of an earlier study [2] in which it was shown that in bean seedlings characterized by certain morphological variations from type, the development of primordial leaf tissue is less than in normal controls grown under conditions as nearly as possible identical. The data then available indicated that a reduction of the volume of primordial leaf tissue is associated with abnormalities of all the types studied, but that the type of variation influences, in some degree, the amount of reduction.

In these first experiments the conclusions were based on primordial leaves only.

The use of such leaves has the obvious disadvantage that they are completely formed in the seed, and undergo merely an enormous expansion (and an undetermined amount of differentiation) in the

---

[1] Studies on the Correlation between Morphological and Physiological Characters, V. Studies I–IV of the series are to be found in *Genetics* 1: 185–196. 1916; 2: 186–212. 1917; 2: 282–290. 1917.

[2] Harris, J. Arthur. Studies on the Correlation of Morphological and Physiological Characters: The Development of the Primordial Leaves in Teratological Bean Seedlings. *Genetics* 1: 185–196. 1916.

germination of the seed and the development of the plantlet to the stage at which measurements were made.

Since the development of the primordial leaves during the germination and establishment of the seedling is relatively great, it seemed quite legitimate to use the weight of green tissue produced by these leaves as a measure of the physiological capacity of seedlings of various types. The fact that these leaves are differentiated in the seed, does, however, constitute a valid objection against their use as a measure of the physiological capacity of the seedling. For such purposes a constant based upon some organ developed later in the life of the individual is desirable.

One of the purposes of this paper is to present the results of determinations upon a later developed organ. The one chosen is the first trifoliate leaf.

This leaf was used because groups of plants of a higher degree of uniformity can be selected at the time of maturity of this leaf than at any later stage in the development of the plant, and because the first compound leaf reaches a degree of maturity sufficient for the purposes of the present study before the primordial leaves are too old to be used for a series of determinations. It is, therefore, possible to repeat, at a slightly later stage of development of the plant, the determinations made on the primordial leaves in the first study as a basis of comparison with the work already done and with the series of constants to be obtained for the first compound leaves of the same plants.

In the first investigation the green weight of the leaf tissue served as the fundamental measurement. In addition to this character certain measurements on the sap properties were also made. In the study of the saps some difficulties were encountered, and it seemed most desirable to discontinue that phase of the study temporarily and to carry out determination of dry weight and water content instead. These new measurements have, therefore, been added to these for green weight.

Materials and Methods

The materials upon which this study is based are the same as those previously employed—a mixture of slightly different strains of navy beans. The seeds which were germinated in the fall and winter months of 1916 were grown in field cultures in 1915.

Seeds from individual plants were germinated in sand. In sorting, the morphologically aberrant seedlings were laid aside with a normal plant to serve as a check for each abnormal. An abnormal and a control seedling *from the same seed plant and germinated in the same*

seed flat were potted side by side in a three inch pot and allowed to grow to the proper stage of maturity under conditions as favorable as we were able to give them.

Before the samples were taken, the plants were carefully inspected and all pairs, one member of which had died, had been injured or which showed in its subsequent development any abnormality in addition to these specified were discarded. Note that there was no *direct* selection for the characters of the abnormal plantlets in this process, since both abnormal and control were discarded if *either* was unsuited for the purposes of the experiment.

There probably was a fairly stringent *indirect* selection, since the death rate and the mutilation rate of the variant individuals was probably greater than that of the normals. Thus more pairs were probably discarded because of an injury to or the death of the abnormal member of the pair than because of the death or injury of a normal member.

The probability that the materials were somewhat selected before the physiological measurements discussed in this paper were carried out renders the findings of greater significance than they would otherwise be.

After the pairs of seedlings had grown until the first compound leaf had attained its full size, and the second compound leaf was developing, but before the primordial leaves had materially deteriorated, samples of leaves were taken by nipping off the laminae only, or the laminae and the single petiolule of the terminal leaflet in the case of the compound leaf. These samples of tissues, each from 100 plants, were enclosed in flasks, weighed, and dried to constant weight in a bath surrounded by boiling water.

Thus the technique employed was exceedingly simple. Because of the size of the samples dealt with, the relative infrequency of the abnormalities, and the large number which had to be discarded, the routine has been excessively laborious. For example, the weighings of the 23 samples and checks discussed in the present paper involve 13,800 leaves gathered from 4,600 plants which were secured by germinating and classifying nearly half a million seedlings.

The structural variation in the bean seedling which is probably the simplest, and the most frequent, is a slight vertical separation of the cotyledons which are normally sensibly opposite in insertion. The amount of the separation is difficult to express quantitatively, since it is in some degree dependent upon the length of the axis. In our studies of seedling variation in Phaseolus, three grades of separation of the cotyledons have been recognized. The line of demarcation between these grades is a quite arbitrary one. This is also true of

the line between "normal," and "abnormal" as applied to the distinction between plants which have cotyledons inserted on the same level and those which have one of the pair sensibly higher on the axis than the other. "Slightly but distinctly separated," has been the descriptive term used in our classification schedules. The cotyledons range in position from those which are just perceptibly not inserted on the same level to those which are perhaps two or three or four millimeters apart. So imperceptible is the line of distinction between normal and abnormal plants that in the classification of the seedlings frequent discussions arose concerning the normality or abnormality of individual plants.

In the present paper I am considering only the simplest type of abnormality. This course has been followed for two reasons.

First, the proof of the existence of a physiological differentiation associated with a very slight structural variation is of far greater interest than the demonstration of measurable physiological differentiation associated with great morphological variation. Second, other types of abnormality with which I have dealt are so difficult to secure in satisfactorily large series that the number of samples as yet available is not sufficient to justify detailed comparisons between the different types of abnormality. I hope ultimately to be able to meet these difficulties. For the present the one type of structural deviation dealt with serves to illustrate the method and one phase of the results of the investigations.

PRESENTATION OF DATA

Consider first of all the green weight of the organs selected.

The average green weight of the primordial and of the first compound leaves for plants which are normal except for slight separation of their cotyledons is shown in Table I.

With one single exception, the average weight of the primordial leaves of the normal plants is higher than that of the abnormal plants. In the single exception to the rule, the difference is small in amount. The average weight of the first compound leaf produced by abnormal plants of this class is in every case but one lower than the weight produced by the sensibly normal individuals. The exception to the rule is the same sample as in the case of the primordial leaves.

The average weight of primordial leaf tissue in the abnormal plants is .5873, the average weight for normal plants is .6680, and the average difference −.0807. The differences in mean weights range in the individual samples from +.0074 to −.1286. For the first compound leaf of the same plants the average weight of the tissues from abnormal individuals is .4797, from a normal plant it is .5610, while the average

difference between the sample and the control is −.0813.   The differ-
ences in average weight vary from +.0368 to −.2492.

TABLE I

Mean Green Weight per Plant of Primordial Leaves and of First Compound Leaf

Sample	Primordial Leaves				First Compound Leaf			
	Abnormal	Control	Difference	Percentage Difference	Abnormal	Control	Difference	Percentage Difference
32	.6034	.7096	−.1062	15.0	.5132	.5929	−.0797	13.4
35	.5648	.6767	−.1119	16.5	.5444	.6188	−.0744	12.0
36	.5951	.6361	−.0410	6.4	.5931	.6254	−.0323	5.2
39	.5619	.6277	−.0658	10.5	.5160	.5549	−.0389	7.0
40	.6096	.7052	−.0956	13.6	.5179	.6138	−.0959	15.6
41	.6068	.7304	−.1236	16.9	.4877	.6140	−.1263	20.6
42	.5879	.6141	−.0262	4.3	.4712	.7204	−.2492	34.6
43	.6222	.7508	−.1286	17.1	.5008	.6115	−.1107	18.1
46	.5956	.7160	−.1204	16.8	.4645	.6019	−.1374	22.8
47	.7058	.6984	+.0074	1.1	.5841	.5473	+.0368	6.7
48	.6389	.7272	−.0883	12.1	.5593	.6395	−.0802	12.5
49	.5902	.6674	−.0772	11.6	.4960	.5851	−.0891	15.2
53	.5402	.5990	−.0588	9.8	.4491	.4948	−.0457	9.2
54	.5720	.6530	−.0810•	12.4	.4091	.4547	−.0456	10.0
56	.5380	.5921	−.0541	9.1	.3994	.4646	−.0652	14.0
61	.5193	.5827	−.0634	10.9	.4443	.4811	−.0368	7.6
64	.5853	.7052	−.1199	17.0	.4530	.5848	−.1318	22.5
65	.5747	.6938	−.1191	17.2	.4402	.5717	−.1315	23.0
66	.5886	.6790	−.0904	13.3	.5246	.5960	−.0714	12.0
70	.6853	.7066	−.0213	3.0	.4794	.4998	−.0204	4.1
71	.5639	.6059	−.0420	6.9	.4132	.4534	−.0402	8.9
72	.5565	.6744	−.1179	17.5	.3799	.4882	−.1083	22.2
73	.5033	.6140	−.1107	18.0	.3933	.4887	−.0954	19.5

If these differences be reduced to percentages by using the weight
of the normal plants as a base, as shown in the final columns of each
section of the tables, it appears that the primordial leaves of the
morphologically aberrant plants are from 3.0 to 18.0 percent lighter
than the leaves of the normal plants in the 22 samples in which this
relationship between the two types of plants holds for the primordial
leaves.   Thus the percentages are highly variable.   The average for
the 23 determinations is 11.95 percent.   In the case of the first com-
pound leaves, the percentage reduction ranges from 4.1 to 34.6 with
an average of 14.06 in the 23 samples.   Note that the percentage
shows that the difference between the abnormal and the control sample
is far less in the case of the single exception, sample 47, than it is in
the average series.   Thus it is only 1.1 as compared with an average
value of 11.95 for the primordial leaves and only 6.7 as compared
with the average of 14.06 percent in the compound leaves.

I now turn to a consideration of dry weight.

The primordial leaves of the abnormal plants in which the two cotyledons are slightly separated are, as shown in Table II, lighter

TABLE II

Mean Dry Weight per Plant of Primordial Leaves and of First Compound Leaf

Sample	Primordial Leaves				First Compound Leaf			
	Abnormal	Control	Difference	Percentage Difference	Abnormal	Control	Difference	Percentage Difference
32	.0445	.0537	−.0092	17.1	.0442	.0517	−.0075	14.5
35	.0366	.0483	−.0117	24.2	.0465	.0530	−.0065	12.3
36	.0422	.0457	−.0035	7.7	.0476	.0499	−.0023	4.6
39	.0409	.0467	−.0058	12.4	.0430	.0470	−.0040	8.5
40	.0438	.0511	−.0073	14.3	.0415	.0496	−.0081	16.3
41	.0431	.0526	−.0095	18.1	.0406	.0494	−.0088	17.8
42	.0416	.0504	−.0088	17.5	.0383	.0519	−.0136	26.2
43	.0429	.0532	−.0103	19.4	.0391	.0493	−.0102	20.7
46	.0408	.0501	−.0093	18.6	.0400	.0492	−.0092	18.7
47	.0442	.0446	−.0004	.9	.0442	.0433	+.0009	2.1
48	.0420	.0464	−.0044	9.5	.0444	.0525	−.0081	15.4
49	.0381	.0436	−.0055	12.6	.0397	.0472	−.0075	15.9
53	.0365	.0410	−.0045	11.0	.0399	.0427	−.0028	6.6
54	.0384	.0445	−.0061	13.7	.0339	.0412	−.0073	17.7
56	.0349	.0491	−.0142	28.9	.0331	.0395	−.0064	16.2
61	.0356	.0402	−.0046	11.4	.0383	.0417	−.0034	8.2
64	.0354	.0438	−.0084	19.2	.0341	.0435	−.0094	21.6
65	.0357	.0410	−.0053	12.9	.0344	.0398	−.0054	13.6
66	.0354	.0395	−.0041	10.4	.0381	.0439	−.0058	13.2
70	.0426	.0465	−.0039	8.4	.0407	.0438	−.0031	7.1
71	.0279	.0303	−.0024	7.9	.0274	.0299	−.0025	8.4
72	.0273	.0407	−.0134	32.9	.0265	.0392	−.0127	3.2
73	.0298	.0378	−.0080	21.2	.0315	.0408	−.0093	22.8

than those of the normal controls in every instance. The average dry weight of the abnormal is .0382, that of the control .0452 and the average difference is −.0070 grams. If the differences be expressed as a percent of the control value as a base, they range from less than 1 to nearly 33 percent, with a general average of 15.21 percent.

The results for the first compound leaf are very similar. In 22 of the 23 cases the primordial leaves of normal plants yield a greater weight of dry substance than those of the abnormal plants. The exception to the rule is again sample 47. The average dry weight of the first compound leaf of abnormal plants is .0385, that of normal plants is .0452 and the average difference is −.0067 grams. If the differences be expressed as percentages of the control constants they are seen to range from 3.2 to 26.2, for the 22 series in which the abnormal plants produce a smaller amount of dry substance. The difference in the single exceptional series is small, only 2.1 percent as compared with a general average of 13.36 percent in the 23 samples.

Having shown that the abnormal plants produce both a smaller green weight and a smaller dry weight in both the primordial and in the first compound leaves, the problem of the relative quantities of water and dry materials in the leaves of the two types of plants naturally presents itself for consideration.

The results have been expressed in terms of the percentage of dry substance in the leaves, *i. e.* (dry weight × 100)/green weight.   The constants appear in Table III.

TABLE III

Percent of Dry Matter in Primordial Leaves and in First Compound Leaf

Sample	Primordial Leaves			First Compound Leaf		
	Abnormal	Control	Difference	Abnormal	Control	Difference
32	7.374	7.567	− .193	8.612	8.703	− .091
35	6.480	7.137	− .657	8.541	8.564	− .023
36	7.091	7.184	− .093	8.025	7.978	+ .047
39	7.278	7.439	− .161	8.333	8.469	− .136
40	7.185	7.246	− .061	8.013	8.080	− .067
41	7.103	7.201	− .098	8.324	8.045	+ .279
42	7.076	8.207	−1.131	8.128	7.204	+ .924
43	6.894	7.085	− .191	7.807	8.062	− .255
46	6.850	6.997	− .147	8.611	8.174	+ .437
47	6.262	6.386	− .124	7.567	7.912	− .345
48	6.574	6.381	+ .193	7.938	8.210	− .272
49	6.455	6.533	− .078	8.004	8.067	− .063
53	6.757	6.845	− .088	8.884	8.629	+ .255
54	6.713	6.815	− .102	8.286	9.061	− .775
56	6.487	8.293	−1.806	8.287	8.502	− .215
61	6.855	6.899	− .044	8.620	8.668	− .048
64	6.048	6.211	− .163	7.528	7.438	+ .090
65	6.212	5.909	+ .303	7.815	6.962	+ .853
66	6.014	5.817	+ .197	7.263	7.366	− .103
70	6.216	6.581	− .365	8.490	8.764	− .274
71	4.948	5.001	− .053	6.631	6.595	+ .036
72	4.906	6.035	−1.129	6.976	8.029	−1.053
73	5.921	6.156	− .235	8.009	8.349	− .340

The results are not so consistent as those for the absolute values, green weight and dry weight. This condition is to be expected for two reasons.   First, the abnormal plants show lower values of both green weight and dry weight than the normal controls.   One cannot, therefore, expect such large differences in the indices calculated from these constants as if both measures did not differ in the same direction between abnormal and control series.   Second, two sets of technical operations are involved in the indices, only one in each of the constants used in calculating these ratios.   While every effort to avoid error was made, the probabilities of error in an index are clearly twice as great as in either of the constants upon which it is based.

Notwithstanding these two sources of difficulty in basing conclusions on relative amount of dry substance, there seem clear evidences that the abnormal plants produce relatively as well as absolutely less dry matter than the normals.

In the case of the primordial leaves, there are 20 samples in which the relative dry weight is lower in the abnormal plants as against 3 in which it is higher. In the first compound leaf there are 15 samples in which the relative weight in the abnormal plants is lower, as compared with 8 in which it is higher than the normals.

The average percentage content of dry substance in the primordial leaves of the abnormal seedlings is 6.509 as compared with 6.779 in the normal controls, or a difference of -0.270. The average percent of dry matter in the first compound leaf is 8.030 in the abnormal as compared with 8.080 in the normal, or a difference of $-.050$ percent.

Concluding Remarks

The constants recorded in this paper are the results of one of the phases of an attempt to determine the nature of the relationship between morphological and physiological variations in plants.

The results of the criteria applied are beautifully clear and consistent.

Seedlings of *Phaseolus* which show one of the smallest definite structural variations, the slight vertical separation of the two cotyledons in their insertion on the axis, are differentiated from the structurally apparently normal individuals in their physiological as well as in their morphological characteristics.

This is shown by the facts that the morphologically abnormal plants produce a smaller weight of green leaf tissue, a smaller actual weight of dry substance in the leaf tissue, and a smaller relative weight of dry substance. This is true for both the primordial leaves and the first trifoliate leaf.

OTE ON THE RELATION OF BLOOD FAT TO SEX, AND ON THE CORRELATION BETWEEN BLOOD FAT AND EGG PRODUCTION IN THE DOMESTIC FOWL

BY

OSCAR RIDDLE AND J. ARTHUR HARRIS

NOTE ON THE RELATION OF BLOOD FAT TO SEX, AND ON THE CORRELATION BETWEEN BLOOD FAT AND EGG PRODUCTION IN THE DOMESTIC FOWL.

By OSCAR RIDDLE and J. ARTHUR HARRIS.

(From the Station for Experimental Evolution, Carnegie Institution of Washington, Cold Spring Harbor, New York.)

In a recent number of this *Journal* Warner and Edmond[1] have presented a considerable series of determinations of the blood fat of White Leghorn hens and cocks. Their interesting discussion of this subject requires a few words of comment on our part.

In the treatment of their data these authors have not sufficiently pointed out the bearing of their results upon earlier work,—indeed most of the latter is ignored; the full value of their own work is not clearly brought out; and some conclusions unwarranted by their data are drawn. The purpose of the paper is stated to be "to show the relationship of blood fat of fowls to (1) egg production, (2) presence of food in the alimentary tract, (3) color of legs, etc., and (4) sex."

The clear and outstanding fact to be found in the 94 blood fat determinations made by Warner and Edmond is that *the blood of the actively laying hen contains a very disproportionately large amount of fat; that of the non-laying hen a very disproportionately small amount of fat.* This fact, however, (and others soon to be mentioned) for the blood serum of the fowl was published by Lawrence and Riddle[2] 2 months prior to the time that the samples for the above mentioned 94 blood analyses were drawn for examination. This earlier work of Lawrence and Riddle on the blood of

[1] Warner, D. E., and Edmond, H. D., *J. Biol. Chem.*, 1917, xxxi, 281.
[2] Lawrence, J. V., and Riddle, O., *Am. J. Physiol.*, 1916, xli, 430.

Riddle utilized very much larger quantities of blood and probably the error involved in the individual determinations was thus reduced. The later and larger series of blood fat determinations therefore confirms the results obtained by Lawrence and Riddle.

Most readers of Warner and Edmond's paper will certainly get the impression that the blood of 1 year old hens has an average fat value of 0.407 per cent, while that of 3 year old hens has a value of only 0.171 per cent. The actual relation of these two groups is as 0.199 to 0.171. The group of 3 year old hens which gave an average of 0.171 were all non-layers and only non-layers among the 1 year old hens may properly be compared with them.

It may next be pointed out that the data "on the presence of food in the alimentary tract" do not justify the conclusion drawn nor indeed any conclusion whatsoever. The following statement is made:[7] "It has been reported by Mathews that in man and animals the blood is much richer in fat after they have been eating than it is after fasting. This is not true with hens, as shown from Table VIII. . . . . The [small] difference . . . would indicate that fasting for 16 hours has no effect upon the amount of fat in the blood." In reality, 16 hours is not a fasting period in the fowl. It more nearly represents the length of time between meals, at least on many of the very short days of winter; and certainly there is some food still in the alimentary tract of a fowl for more than 16 hours after feeding.

There are, however, two reasons why the data cannot answer the question of the relative amount of fat in the blood soon after as compared with 16 hours after a meal. These are, first, that the *female* fowl was used, and in these birds the individual variation is enormous, the really decisive factor associated with blood fat values being the "laying" or "non-laying" condition of the hen. Other figures by these authors show that the blood fat of a laying hen (1.953) may be more than twenty times greater than that of a non-laying hen (0.083). Their figures (Table VI) also indicate that this value in laying hens may vary between 0.2 and 1.953. It is therefore obvious that one may not expect to find *differences due to fasting for 16 hours* in the averages obtained

[7] Warner and Edmond, *ibid.*, 289.

f m twelve individuals of one group when compared with twelve
i ividuals of the other group. An illustration of the unsuit-
aility of the data for this purpose is the following: If the twelfth
(st) bird in each series be omitted from the averages, these latter
a changed from 0.405 and 0.396 to 0.268 and 0.263 respectively.
(ly in males, in which the fluctuations of the blood fat values
a much decreased, could one hope for success in a study of this
cestion in the fowl. Still another reason for the inadmissibility
c the data as submitted is that no basis for the selection of the
rticular individuals chosen and compared is given. As the
cta stand, and if one is permitted to select at will the "non-
f ting" birds from the whole group of fat determinations, it is
ssible to have those birds show *more* or show *less* blood fat than
t "fasting" birds. These data, therefore, supply no evidence
t t the fowl is an exception to the well known rule that the per
c t of fat in the blood increases soon after the ingestion of food.
I all of the fat determinations made by Lawrence and Riddle
t s fact was recognized, and the blood samples were all drawn at
aproximately the same time of day; namely, in the early after-
on.

Our next point concerns the correlation between the fat content
c the blood and egg production. Admitting that birds which are
l ing at the time the blood samples are taken differ from those
ich are not laying, a further problem arises concerning the
 ationship of the percentage of fat found in the blood and the
t al egg record of the bird for the year.

These authors say:[8] "The fat content of the blood is corre-
l ed with egg laying activity, and *there is also a slight correla-
t n between the amount of fat in the blood and high yearly production*."
But in their summary, they conclude:[9] "*There is little or no
crelation between the amount of fat in the hen's blood and her yearly
e yield*. On the other hand the blood of a hen laying at the
tne the sample is taken is much richer in fat than that of a hen
ich is not laying."[10]

Now both of these contradictory statements cannot be true.
 ere is, indeed, no reason why statements concerning correlation

[8] Warner and Edmond, *ibid.*, 288.
[9] Warner and Edmond, *ibid.*, 293.
[10] The italics are ours in both quotations.

should be of a general and vague nature. Correlation is a very definite quantitative phenomenon, measurable on a universally applicable scale of -1 to $+1$. The numerical value of the correlation coefficient, r, may be determined in any instance for two measured variables, such for example, as fat content and annual egg record, by very simple and well known formulas.

We now turn to the actual results which may be obtained by applying the correlation formulas[11] to the data published by Warner and Edmond. Using a simple method of direct summation of the actual values, their squares and their products,[12] we find the following results for the whole series of 1 year old hens:[13]

$$r_{fe} = +0.247 \pm 0.076$$

Thus, taking the whole series of 1 year old birds examined, the hens which have a larger amount of fat in their blood in October have, on the average, laid a larger number of eggs during the year

[11] Warner and Edmond must have been fully acquainted with the advantages of applying the correlation formulas to such problems as these for the data contained in the preliminary papers by Blakeslee and Warner cited and discussed by Warner and Edmond, together with far more extensive data collected since these two preliminary papers were published, have been carefully treated in great detail by the statistical methods in a paper by Harris, Blakeslee, and Warner (*Genetics*, 1917, ii, 36; *Proc. Nat. Acad. Sc.*, 1917, iii, 237). These papers, like that by Lawrence and Riddle, they have, quite inadvertently, failed to cite, although they contain much that throws light upon the problems considered by them. A further investigation of the problem of the interrelationship between egg laying activity at various periods which also contains materials bearing very directly upon the problem of the physiology of egg production is now in press (*Genetics*, 1918, iii, 27).

[12] Harris, J. A., *Am. Naturalist*, 1910, xliv, 693.

[13] Those unfamiliar with statistical formulas need only remember that the correlation coefficient measures the intensity of interrelationship between two variables on the scale of -1 to $+1$. Thus if annual egg production is lower in birds with lower percentages of fat and higher in birds with higher percentages of fat correlation between egg production and fat content is positive in sign and lies somewhere between no correlation and perfect correlation. If high annual egg record is associated with low fat content and low annual record with high fat content, correlation is negative in sign and is measured by a coefficient lying between 0 and perfect negative correlation.

But this group is highly heterogeneous. It comprises birds which were still laying at the time the blood samples were taken and those which had ceased to lay.[14] Dividing the birds into two groups on the basis of their laying activity at the time of sampling, we find the following values for averages, variabilities, and correlations.

For 54 birds which had ceased laying,

$$\bar{f} = 0.199 \qquad \bar{e} = 139.09$$
$$\sigma_f = 0.0847 \qquad \sigma_e = 38.948$$
$$V_f = 42.36 \qquad V_e = 28.00$$
$$r_{fe} = -0.296 \pm 0.084$$

For 16 birds which were still laying at the time the samples were taken,

$$\bar{f} = 1.103 \qquad \bar{e} = 163.37$$
$$\sigma_f = 0.6102 \qquad \sigma_e = 30.503$$
$$V_f = 55.30 \qquad V_e = 18.67$$
$$r_{fe} = +0.351 \pm 0.147$$

Here f = percentage fat, e = number of eggs laid during the year, the bars over these letters denote the averages, the sigmas their absolute variabilities in terms of square root of mean square deviation from their mean, V their coefficients of variation, i.e., 100σ / mean, and r the correlation between them.

Now while the probable errors of these correlations are large, because of the relatively few determinations, the correlations suggest at once that *while the birds are laying* those which have the highest percentages of blood fat are those which make the highest egg records, but that *after the bird has ceased laying in the autmn* those which have laid the greatest number of eggs during the year have the blood most depleted of fat; or, at any rate, have blood with least fat.

Support for this view is furnished by splitting the birds up into two groups according to the color of the beak, legs, and vent— a character which when taken in October may be taken as a criterion of annual egg production.[15]

[14] One bird had not laid at all during the year.
[15] Blakeslee and Warner, as cited by Warner and Edmonds. See also Harris, Blakeslee, and Warner, *Genetics*, 1917, ii, 36; *Proc. Nat. Acad. Sc.*, 1917, iii, 237. Blakeslee, Harris, Warner, and Kirkpatrick, *Storrs Agric. Exp. Station*, *Bull. 92*, 1917.

Birds with yellow beaks, legs, and vent, N = 23
$$r_{fe} = 0.411 \pm - 0.117^*$$

Birds with pale beaks, legs, and vent, N = 18
$$r_{fe} = + 0.532 \pm 0.114$$

* This is the value obtained by using only the twenty-three 1 year
birds included in Table VII of Warner and Edmond. If the nine 3 year
birds which they have lumped with these are included the correlatio
$r = - 0.332 \pm 0.106$. The correlation for the nine 3 year old birds al
is $r = - 0.021 \pm 0.225$.

The results are the same in sign but numerically larger th
the coefficients based on the actual record concerning laying
tivity. They point to the same conclusion; namely, that w
birds are in the actively laying condition there is a *positive*
relation between annual egg record and per cent of fat in the blo
but that *after the egg-producing capacity of the bird has been
hausted*, there is a *negative* correlation between egg record and
amount of fat in the blood.

If this is true, one might expect the correlation to change
the lapse of time since laying. This point may be tested
splitting the data of Warner and Edmond into groups accor
to the number of days since laying to the time the blood san
was taken, and determining the correlation between fat con
and egg production in each of these groups separately. We fir

For birds laying at time blood sample was taken, N = 16
$$r_{fe} = + 0.351 \pm 0.148$$

For birds which had not laid for from 1 to 24 days, N = 16
$$r_{fe} = + 0.054 \pm 0.168$$

For birds which had not laid for from 25 to 29 days, N = 22
$$r_{fe} = - 0.132 \pm 0.141$$

For birds which had not laid for from 60 to 365 days, N =
$$r_{fe} = - 0.620 \pm 0.103$$

Thus, as suggested above, there is a progressive change in
nature of the correlation between fat content and total egg
cords which is positive for birds in a laying condition, sink
zero after the cessation of laying, and finally takes a high neg
value in birds which have long since ceased to lay.

ummarizing these statements on correlation we note: At the en of the first laying year, in October, the correlation between th per cent of fat in the blood and the annual egg record of the yr is positive in the case of birds which are still laying, but ntative in the case of birds which have ceased laying; or, in oter words, birds which have laid larger numbers of eggs and are stl laying have a higher percentage of fat in their blood than lay- in birds which have made a poor record during the year. Birds wch have laid a large number of eggs, but have exhausted their fe ility, have a smaller per cent of fat in their blood than non- laing birds which have made poor egg records.

By dividing records into classes, according to the number of drs since the cessation of laying, the correlation between October bod fat content and annual egg record changes systematically frn a positive to a negative relationship.

This brief discussion will perhaps indicate how useless it is to dw conclusions concerning correlations between blood fat and anual egg production without actually determining the cor- ration coefficients. The constants given here are of course suject to the limitations of the series of data upon which they are bed. Larger sets of determinations, made at various times of th year, when analyzed by proper statistical formulas, should yld further important information on the relation of fat metab- om to the functioning of the ovary.

Addendum.—In the foregoing treatment of the data published by Warner al Edmond[1] we had two purposes: (a) to call the attention of biological clmists to earlier work on the same subject which those authors entirely igored, and (b) to clear up some matters concerning the interpretation othe data presented by them. We regret, therefore, that their second cmmunication[16] not only fails to contribute to these desirable ends but irroduces other errors, some misleading quotations and statements, and mch wholly irrelevant discussion.

It therefore seems quite useless, as well as personally undesirable to us, tcontinue this discussion by commenting further on their second publica- tin. For the convenience of the reader who may, if he cares to take the tie, verify the statements from the original papers, we may add this sum- nry.

1. A note by Warner and Edmond in a poultry publication appeared a mnth after the work of Lawrence and Riddle was published. It dealt

---

Warner and Edmond, *J. Biol. Chem.*, 1918, xxxiv, 171.

with methods of selecting high producing and low producing hens, a
their only reference to the problem of blood fat in the fowl has be
fully quoted by us.[3] Warner and Edmond published nothing on blood
in relation to *sex* nor in relation to *laying* and *non-laying* until 11 mon
after the paper by Lawrence and Riddle appeared, which dealt exclusiv
with these problems and not with that of high and low production. Th
samples for this purpose were not drawn until 2 months after this publi
tion.

2. The data of Warner and Edmond when properly classified and a
lyzed confirm the conclusions of Lawrence and Riddle in so far as the pr
lem of sex and of laying and the non-laying condition of the bird is c
cerned.

3. The data of Warner and Edmond are not sufficiently extensive
decide finally the question of the possible relationship between percenta
of blood fat and high and low production. When properly classified b
logically and adequately analyzed statistically they indicate the inc
rectness of the conclusion of Warner and Edmond[17] that, "There is little
no correlation between the amount of fat in the hen's blood and her yea
egg yield."

---

[17] Warner and Edmond, *J. Biol. Chem.*, 1917, xxxi, 293.

THE EGG RECORDS OF LIMITED PERIODS AS CRITERIA FOR PREDICTING THE EGG-PRODUCTION OF THE WHITE LEGHORN FOWL

J. ARTHUR HARRIS, WM. F. KIRKPATRICK, A. F. BLAKESLEE,
D. E. WARNER AND L. E. CARD

Station for Experimental Evolution, Cold Spring Harbor, Long Island, New York
and
Storrs Agricultural Experiment Station, Storrs, Connecticut

From the purely scientific side the interrelationship of the egg production of different periods is a problem of great biological interest. From the economic side, two ends which may be either quite distinct or interdependent, are to be attained by the development and application of formulae for the prediction of the egg production of a bird during any period of the pullet year. The first is the determination of the probable future record of an individual bird from her past performance, as a basis for the decision as to whether she shall be kept for egg production or sold for meat. The second is the estimation of the annual record of a bird as a basis of decision as to whether or not she shall be kept until the following season to be used as a breeder.

It will be evident to those who have had to consider the problems with which we have to deal, that economic factors,—particularly the cost of trap-nesting,—and the purpose for which prediction is being made will have great weight in determining the period and the number of periods to be used in the prediction equations. In determining which birds shall be sold to the commission man and which may be fed with reasonable prospects of profitable returns for the remainder of the year, the breeder is not concerned primarily with the record which the bird makes for the year as a whole. Practically he requires to know what returns she will make for the remainder of the period over which she may be retained in the flock. He has already maintained her for n of the 12 months. The question which the poultryman would like to have answered is whether her record during this period has been such that he can afford to feed and house her for the remaining $12 - n$ months. It is evident that to be of the greatest value for this purpose the prediction should be made from periods as early as possible in the life of the bird. In other words, if birds are to be culled out of the flock and sold for their flesh because they are unprofitable as egg producers, this should be done at a time when the maximum saving in cost of maintenance can be coupled as closely as possible with the maximum sale price.

The correlations required for this purpose are, therefore, those between the record of any period which may be selected as a basis for judgment, and the record of later months.

Since in the selection of birds to be held over for breeders the total annual production is presumably the factor to be chiefly taken into consideration it is evident that the correlations to be determined are those between the record of the individual months and of the year as a whole. It would, of course, be better for this purpose if the records of the entire year were known, but as pointed out by CARD (1917, p. 66) many poultry-

:en who feel that they cannot trap-nest for the whole year may be able to
o so for a period of one, two or three months. Furthermore, another
:ethod of determining the record of individual birds for a given period
hen a knowledge of the maternity of the individual eggs is not necessary
is recently been suggested by ALDER and EGBERT (1918).

In selecting the months to be used in the calculation of prediction equa-
ons for practical use one should of course be guided by the experience
.id needs of commercial poultrymen. It has generally been maintained
iat culling should not be undertaken until the period of summer pro-
uction. The price of eggs is highest during the winter and early spring
ionths. If birds have been left in the flock during these months in the
ope that they will produce some eggs in the period of generally low pro-
uction and high price, it would be unwise to discard them just at the time
hen some return will most probably be made for maintenance rations
uring a preceding period, unless the price of eggs is too low to pay for
ie cost of the feed during the period of spring production when most of
ie birds are laying heavily. Possibly this opinion is partly based on the
iadequacy of criteria for eliminating the probably less productive birds
arly in the season.

Furthermore, the physiological laws underlying egg production should,
i so far as possible, be taken into account in selecting the periods upon
hich prediction equations are to be based. Previous students of fecun-
ity in the domestic fowl have been inclined to recognize certain standard
eriods. For example, the division of the year adopted in the many papers
ublished by PEARL, SURFACE and their collaborators at the MAINE STA-
ION is the following:

Period of winter egg production: November, December, January and
'ebruary.

Period of spring egg production: March, April and May.

Period of summer egg production: June, July and August.

Period of autumn egg production: September and October.

This is also the division adopted by CARD (1917, p. 67).

These periods will be used in the present paper, and other combinations
dded so that there may be a wide range of periods from which selection
iay be made.

While fully recognizing the foregoing principles we have felt it a part
f the biologist's task to furnish a comprehensive series of measures of
iterrelationships of the egg production of various periods from which
iny desired equation may be selected. It has, therefore, seemed to us
iest in this publication to base prediction equations upon a large number

of months and combinations of months in order (a) to determine the months which give the best results and (b) to enable those who wish to predict from any group of months.

In the investigations one phase of which is presented in this paper, we have sought among other things:

(1) To determine the best method of predicting the annual egg production of a bird from the known record of any individual month.

(2) To determine the best method of predicting the annual egg production of a bird from the combined records of two or more months.

(3) To determine the best method of predicting the egg record of a bird during a portion of the year from the record of a single antecedent month or a group of antecedent months.

(4) To compare the relative merits of these several methods of prediction among themselves and to determine thereby which of the methods make possible the most exact prediction as a basis for determining which is likely to be of the greatest practical value.

We fully recognize, and desire to emphasize especially, the fact that the whole problem of the prediction of future egg production cannot be solved in a single investigation. The problem is exceedingly complex and number of factors are not taken into consideration at all in the present paper. All that has been attempted is to indicate the possibility of a important line of advance and to lay the foundations, in a series of statistical constants, for wider investigations. Some of these are already in progress. In the meantime, the results presented here may prove useful both from the practical standpoint and in facilitating to some extent further and more adequate investigations.

The first definite step in the direction of the use of the egg record of short recorded period for the prediction of the probable production during a subsequent or a longer period was, as far as we are aware, taken in 191? when it was shown (HARRIS, BLAKESLEE, WARNER and KIRKPATRICK 1917) that in a heterogeneous series of birds such as are submitted by practical breeders in egg-laying contests, the October egg production correlated with that of every other month of the year. The whole subject was carried much further in a second memoir (HARRIS, BLAKESLEE and KIRKPATRICK 1917, 1918) in which the correlations between the records of the individual months and the production of the whole year, between the records of the individual months and of the remaining 11 months of the year, and between the production of 5 of the individual months and the production of all the other individual months, were published for two series of birds. In this paper the equations for the prediction of total

annual production from the record of the individual months were given.

The results given in our second paper (HARRIS, BLAKESLEE and KIRK-PATRICK 1917, 1918) show clearly that it is possible to predict with a considerable degree of accuracy the annual egg production of a group of birds from their record for a given month. They also indicate that it is possible within limits to predict the egg production of any month, p, from the egg record of any other month, q.

Almost simultaneously CARD (1917) considered the correlation between the records of various periods as a basis for the prediction of annual egg production. Prediction equations were not, however, given.

While the determination of equations for the prediction of the egg production of a subsequent or a more extended period from the actually recorded production of a limited period must rest upon biometric theory, we have deemed it proper in the testing of these equations to proceed in a purely objective manner.

We have determined a series of prediction equations and have used these equations for estimating the egg production of a series of birds, the egg record of which is unknown as far as the development of the equations is concerned. We then determined the difference between the yield predicted by the equations and the actual yield in the case of each individual bird. The average of these deviations, or any other suitable mathematical constant based upon them, furnishes a criterion of the suitability of the equation for purposes of prediction. That equation is best which predicts most exactly the annual egg yield, or the egg production of any shorter period, for a bird of which the record of a limited period is known.

Since the birds entered in the INTERNATIONAL EGG-LAYING CONTEST at Storrs are drawn from a wide geographical area and are furnished by a large number of breeders, and since the conditions in the different years are maintained as nearly constant as possible, it seemed desirable to utilize records from this contest subsequent to those upon which the equations are based in testing the value of the equations. The problem is: How closely can the actual production of a bird entered in the contest in a given year be predicted from equations based on the records of previous years when one or more months' performance of this bird is known from observation? We have, therefore, as already noted, based the test of our series of equations first of all upon the records secured in Connecticut during the contest year 1917 and 1918.

The equations which we publish are based upon 1840 single-comb White Leghorn birds entered in the INTERNATIONAL EGG-LAYING CONTEST for

the years 1911 to 1917.   The prediction equations have been tested upon
415 birds whose records were obtained during the year extending from
November 1, 1917, to October 31, 1918.

The justification for the course followed is found in the general principle
that a theory should not be tested against the observations upon which
it is based.

For practical reasons this paper is limited to a test of the accuracy with
which the egg record of a series of 415 birds trap-nested at Storrs during
1917–1918 can be predicted by a series of linear equations based on the
experience of the six preceding years, 1911–1917, at the same place.   It
may be urged that conditions at Storrs are not representative of those
prevailing in different parts of the country.   Recognizing, for the sake of
argument at least, the validity of this objection we have been glad to avail
ourselves of records taken elsewhere.   These are now being used to test
the accuracy with which the production of birds in any locality may be
predicted by means of equations based primarily upon experience in another
place or with another series of birds.   The results of these studies will
eventually be published.

NOTATION AND THEORY EMPLOYED

We shall find it convenient to have a simple and rigid notation.   Let
e represent the recorded egg production of a bird in any month, Σ denote
a summation of monthly egg records for a given bird, 1, 2, 3, . . . 12
denote the twelve successive months of the pullet year, i.e., the November
of the year in which the bird was hatched until and including the following
October.   Then $e_1, e_2, e_3, \ldots, e_{12}$ represent the November, December,
January, . . ., October egg record of a bird with an annual record of
$E = \overset{12}{\underset{1}{\Sigma}} (e)$ eggs.   Further, E_n denotes the total number of eggs laid in
any month or group of months subsequent to any given month or group
of months used as a basis of prediction, i.e.,

$$E_{11} = E - e_1 = \overset{12}{\underset{2}{\Sigma}} (e), \ E_{10} = E - e_1 - e_2 = \overset{12}{\underset{3}{\Sigma}}(e), \ \ldots, E_1 = e_{12}.$$

In the present paper we have used only the linear prediction equation
derived from the means, standard deviations and product-moment coeffi-
cients of correlation between the periods, or groups of periods, of egg pro-
duction, i.e., with equations of the type

$$E = \left(\overline{E} - r_{e_p E} \frac{\sigma_E}{\sigma_{e_p}} \overline{e_p} \right) - r_{e_p E} \frac{\sigma_E}{\sigma_{e_p}} e_p$$

where E represents the annual egg production or the production of any period of months, and e_p denotes the production of any period used as a basis for prediction.

The reader may quite legitimately suggest that in certain cases better prediction might have been secured by the use of regression curves of a higher order. This may be true. Our plan has been to test not merely the linear equations but others as well. Considerable progress has been made toward this end. Comprehensive tests will, we hope, eventually be published. Since, however, a relatively high degree of accuracy of prediction may be attained in most cases by the use of the linear equation, it does not seem proper to withhold useful results until it is possible to determine whether additional refinement can be attained.

The essential characteristics of equations for the prediction of egg yield are two:

1. That the errors of prediction be distributed about the true numbers in such a manner that estimations will not in the long run be either too high or too low.

2. That the magnitude of the deviations of the predicted from the observed egg productions be as small as possible.

Thus in testing formulae by determining how efficiently they predict the production of birds whose record is actually known, we shall consider that formula the best which (a) shows the least error in the direction of consistently too high or too low prediction, and (b) gives the lowest deviation of the predicted from the observed record.

To test the first of these essentials we have merely to determine the average deviation with regard to sign of the predicted from the actually measured egg production. This is given by

$$\frac{\Sigma \, (E'_p - E_p)}{N}$$

where E_p is the actual egg production of a bird, E'_p the theoretical egg record of an individual bird for a period p, and N the number of birds considered. Here a negative sign indicates that the equation has predicted records which are on the average too low, whereas a positive sign indicates that it has predicted records which are on the average too high.

But, as noted above, a formula must do more than fail to consistently overestimate or underestimate. It must give predicted values which show the lowest possible deviation from those determined by trap-nesting. We have, therefore, to consider the test which shall be applied to deter-

mine which formula gives the lowest deviation. Two methods may be suggested.

First, the deviations may be summed without regard to sign and divided by their number. This gives an average deviation without regard to sign of the predicted from the recorded production for any flock and period under consideration.

The disadvantages of this method are two: (a) It ignores mathematical convention with regard to signs. (b) It gives large and small deviations a weight proportional to their actual magnitudes. Thus 50 deviations of 3 eggs each and 50 deviations of 5 eggs each would give an average deviation of 4 eggs, while 50 deviations of 1 egg each, 25 deviations of 6 eggs each and 25 deviations of 8 eggs each would also give a general average deviation of 4 eggs. But since one of the ideals to be attained in the selection of a formula would seem to be to obtain one which will avoid the grosser errors it seems proper to weight the larger deviations. This can be most logically done by squaring. Then

$$\left\{ \frac{\Sigma (E'_f - E_f)^2}{N} \right\}^{\frac{1}{2}}$$

gives a square root of mean square deviation, or a "root mean square deviation." This is probably the best available measure of the deviation of prediction from observation.

For completeness we shall employ all three methods in the tests of equations used in this paper.

The method of taking the difference has been so chosen that a positive sign, indicating larger error of estimating, shows an inferiority in the equation.

Two of the criteria are values without sign. In the case of the average deviation with regard to sign the criteria may be either positive or negative. In comparing two different methods of prediction we have considered that the *magnitude* of the error and not the *sign* is the critical point. In such comparisons, therefore, all of the criteria have been considered as alike in sign. Cases may possibly arise in which it is desirable to consider the question of over prediction or under prediction by two formulae which may be under consideration. If so our tables of criteria and not the differences as published should be consulted by the reader.

The characteristic equation given above is strictly valid only when applied to the population from which it is deduced. Its extension without modification to another population is justified only if the physical constants

TABLE 1

Comparison of constants for series of birds on which equations were based (1911 to 1917) and series on which they were tested (1917-1918).

| | MEAN | | | | | STANDARD DEVIATION | | | | |
	1911 to 1917	1917 to 1918	Difference	$\frac{\text{Diff}}{E_{\text{Diff}}}$	Percentage difference	1911 to 1917	1917 to 1918	Difference	$\frac{\text{Diff}}{E_{\text{Diff}}}$	Percentage difference
November	5.20±0.09	5.78±0.20	+0.58±0.21	2.76	11.11	6.00±0.06	6.34±0.14	+0.34±0.15	2.26	5.66
December	6.58±0.11	6.21±0.22	-0.37±0.24	1.54	5.62	7.10±0.07	6.80±0.15	-0.30±0.16	1.87	4.22
January	6.07±0.10	6.63±0.19	+0.56±0.21	2.66	9.22	6.63±0.07	5.83±0.13	-0.80±0.14	5.71	12.06
February	10.10±0.09	9.93±0.19	-0.17±0.21	0.80	1.68	5.91±0.06	5.94±0.13	+0.03±0.14	0.21	0.50
March	17.45±0.08	16.96±0.18	-0.49±0.19	2.58	2.80	5.31±0.05	5.52±0.13	+0.21±0.13	1.62	3.95
April	18.85±0.07	17.48±0.18	-1.37±0.19	7.21	7.26	4.67±0.05	5.44±0.12	+0.77±0.13	5.92	16.48
May	20.55±0.08	21.84±0.16	+1.29±0.17	7.58	6.27	5.24±0.05	5.06±0.11	-0.18±0.12	1.50	3.43
June	20.41±0.09	19.18±0.20	-1.23±0.21	5.85	6.02	5.84±0.06	6.13±0.14	+0.29±0.15	1.93	4.96
July	19.28±0.10	17.96±0.24	-1.32±0.26	5.07	6.84	6.43±0.07	7.43±0.17	+1.00±0.18	5.55	15.55
August	17.10±0.11	16.71±0.26	-0.39±0.28	1.39	2.28	7.30±0.08	7.96±0.18	+0.66±0.19	3.47	9.04
September	11.78±0.13	12.98±0.26	+1.20±0.29	4.13	10.18	8.36±0.09	7.93±0.18	-0.43±0.20	2.15	5.14
October	4.92±0.10	5.87±0.23	+0.95±0.25	3.80	19.30	6.63±0.07	7.07±0.16	+0.44±0.17	2.58	6.63
Annual	158.36±0.68	157.59±1.41	-0.77±1.56	0.49	0.48	43.34±0.48	42.63±0.99	-0.71±1.10	0.64	1.63

and the correlations of the variables in the two populations are essentially identical.

Because of the uniformity of care and the wide origin of the birds exhibited each year at the INTERNATIONAL EGG-LAYING CONTEST at Storrs the average productions do not differ widely in the different years. Thus the monthly and annual averages and standard deviations for the 1840 birds upon which the equations were based and the 415 birds upon which they were tested appear in table 1.[2]

While certain of the differences are significant in comparison with their probable errors it is quite clear that the averages for the two periods are in fair agreement.

Bird 997, Pen 100

	1	2	3	4	5	6	7	8
N		143.1	−17.9	320.41				
D		137.2	−23.8	566.44	161	143.1	−17.9	320.41
J	3	148.1	−12.9	166.41	161	134.4	−26.6	707.56
F	13	169.7	+8.7	75.69	158	135.6	−22.4	501.76
M	24	189.1	+28.1	789.61	145	136.2	−8.8	77.44
A	21	169.2	+8.2	67.24	121	129.2	+8.2	67.24
M	29	199.6	+38.6	1489.96	100	99.4	−0.6	0.36
J	25	180.4	+19.4	376.36	71	92.3	+21.3	453.69
J	27	193.2	+32.2	1036.84	46	61.8	+15.8	249.64
A	13	142.0	−19.0	361.00	19	44.9	+25.9	670.81
S		118.5	−42.5	1806.25	6	12.8	+6.8	46.24
O	6	162.0	+1.0	1.00	6	−9.7	−6.7	44.89
Year	161							

The method followed in the calculations may be illustrated by one o the calculation blanks for the individual bird—No. 997, pen 100. Th first column shows the production for the month indicated by the letter on the stub. This serves as the basis of prediction. The second colum shows the predicted number of eggs for the year, the third shows th deviation of this predicted number from the annual total of 161 egg The fourth column gives the squares of these deviations of prediction fror observation. The fifth column shows the number of eggs in the remainin months of the year.[3] The sixth column shows the number of eggs predicte

[2] The percentage differences have been calculated by using the monthly averages for 19 to 1917 as a base.

[3] The yields for the remaining months (columns 5 to 8) are dropped one space so as to coi cide with the first month of the period. For example, bird 997 laid 161 eggs in the period fro December to October; 161 in the period from January to October; 158 in the period from Fe ruary to October, and so on.

for the remaining months. The seventh shows the deviations and the eighth the squares of the deviations of these predicted values from the actual record for the remaining months.

Calculation blanks for each individual bird were made on this principle for each of the equations used. The labor of testing the equation has, therefore, been very heavy, involving the calculation of 29,465 predicted values and the summations of the errors and squares of errors of the deviations of these predicted records from their true value.

The excessive arithmetical routine has been ably handled by Miss EDNA M. PECKHAM, Miss IDA M. PECKHAM, Miss RUTH T. CRAWSON, and Miss KATHLEEN GAVIN of the Biometric Laboratory of the STATION FOR EXPERIMENTAL EVOLUTION. We are indebted to Miss EDNA K. LOCKWOOD for the diagrams, as well as for much assistance in the computations.

TESTS OF EQUATIONS EMPLOYED

Prediction of annual production from the record of one month

Consider first of all the results of the attempts to predict the annual egg production of 415 White Leghorn birds observed at Storrs in 1917–1918 from the records of a single month's production. The equations based on the 1911 to 1917 experience are as follows:

Month from which prediction is made	Prediction equation
November	$E = +143.186 + 2.914\, e_1$
December	$E = +137.293 + 3.200\, e_2$
January	$E = +138.271 + 3.308\, e_3$
February	$E = +118.689 + 3.926\, e_4$
March	$E = +76.160 + 4.708\, e_5$
April	$E = +62.688 + 5.074\, e_6$
May	$E = +58.009 + 4.883\, e_7$
June	$E = +59.977 + 4.818\, e_8$
July	$E = +71.137 + 4.523\, e_9$
August	$E = +90.391 + 3.974\, e_{10}$
September	$E = +118.509 + 3.381\, e_{11}$
October	$E = +141.470 + 3.429\, e_{12}$

These are in good general agreement with the equations for two of the years, 1913–1914 and 1914–1915, published in a former paper (HARRIS, BLAKESLEE and KIRKPATRICK 1918, page 33, table 5). The graphical tests for linearity of regression (loc. cit., diagrams 2–5, p. 34–39) for these two years, indicate a fairly close approximation to linearity throughout the greater part of the range of variation of monthly egg production. A critical test of linearity presents some difficulties because of the concen-

tration of the bulk of the birds into a few of the classes, with the result
that a rather large number of classes contain only a few birds each. A
closer study of the fit of the regression line may, therefore, be deferred
until more data are in hand.

The results of the tests of accuracy of prediction in the 415 White Leg
horn birds of the 1917–1918 contest are given in tables 2 to 4. Since later

TABLE 2

Average deviation with regard to sign of predicted annual egg record from actual record. Prediction
of annual production from one- and from two-months performance. Equations based on Storrs
experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM TWO MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFER-ENCE IN PERCENT AGE DEVIATIC
	Base of prediction	Actual deviation	Percent-age deviation	Base of prediction	Actual deviation	Percent-age deviation		
For the whole year	November	+2.39	1.52	Nov. + Dec.	+1.16	0.74	+1.23	+0.78
	December	−0.49	0.31	Nov. + Dec.	+1.16	0.74	−0.67	−0.43
	December	−0.49	0.31	Dec. + Jan.	+1.15	0.73	−0.66	−0.42
	January	+2.58	1.64	Dec. + Jan.	+1.15	0.73	+1.43	+0.91
	January	+2.58	1.64	Jan. + Feb.	+1.75	1.11	+0.83	+0.53
	February	+0.06	0.04	Jan. + Feb.	+1.75	1.11	−1.69	−1.07
	February	+0.06	0.04	Feb. + Mar.	−1.22	0.77	−1.16	−0.73
	March	−1.63	1.03	Feb. + Mar.	−1.22	0.77	+0.41	+0.26
	March	−1.63	1.03	Mar. + Apr.	−4.94	3.13	−3.31	−2.10
	April	−6.23	3.95	Mar. + Apr.	−4.94	3.13	+1.29	+0.82
	April	−6.23	3.95	Apr. + May	+0.48	0.30	+5.75	+3.65
	May	+7.02	4.45	Apr. + May	+0.48	0.30	+6.54	+4.15
	May	+7.02	4.45	May + June	+0.82	0.52	+6.20	+3.93
	June	−5.21	3.31	May + June	+0.82	0.52	+4.39	+2.79
	June	−5.21	3.31	June + July	−6.60	4.19	−1.39	−0.88
	July	−5.27	3.34	June + July	−6.60	4.19	−1.33	−0.8
	July	−5.27	3.34	July + Aug.	−3.81	2.42	+1.46	+0.9
	August	−0.82	0.52	July + Aug.	−3.81	2.42	−2.99	−1.9
	August	−0.82	0.52	Aug. + Sept.	+2.60	1.65	−1.78	−1.1
	September	+4.78	3.03	Aug. + Sept.	+2.60	1.65	+2.18	+1.3
	September	+4.78	3.03	Sept. + Oct.	+5.34	3.39	−0.56	−0.3
	October	+3.95	2.51	Sept. + Oct.	+5.34	3.39	−1.39	−0.8

we shall have to compare the results for prediction from one month
performance with that from two- and from three-months record it has been
desirable to give the results side by side in the same table. The reader
need not, therefore, concern himself with the values for prediction from
two-months production until later. Since the errors of prediction of the
annual record from each individual month must be compared with the
results for prediction from the combined production of two months, the
constants for the single months have been given in duplicate.

The average errors with regard to sign are generally low, that for prediction from November and from January production gives on the average 2.4 eggs too many for the year. For December, February, March and August the prediction is in error by less than 2 eggs. The values predicted from April, May, June, July, September and October records are about 4 to 7 eggs in error.

TABLE 3

Average deviation without regard to sign of predicted annual egg record from actual record. Prediction of annual production from one- and from two-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM TWO MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
	November	29.59	18.78	Nov. + Dec.	28.09	17.82	+1.50	+0.96
	December	29.26	18.57	Nov. + Dec.	28.09	17.82	+1.17	+0.75
	December	29.26	18.57	Dec. + Jan.	27.23	17.28	+2.03	+1.29
	January	30.09	19.09	Dec. + Jan.	27.23	17.28	+2.86	+1.81
	January	30.09	19.09	Jan. + Feb.	27.35	17.35	+2.74	+1.74
	February	27.28	17.31	Jan. + Feb.	27.35	17.35	−0.07	−0.04
	February	27.28	17.31	Feb. + Mar.	25.04	15.89	+2.24	+1.42
	March	27.95	17.73	Feb. + Mar.	25.04	15.89	+2.91	+1.84
	March	27.95	17.73	Mar. + Apr.	26.74	16.97	+1.21	+0.76
For the whole year	April	28.72	18.22	Mar. + Apr.	26.74	16.97	+1.98	+1.25
	April	28.72	18.22	Apr. + May	26.68	16.93	+2.04	+1.29
	May	28.62	18.16	Apr. + May	26.68	16.93	+1.94	+1.23
	May	28.62	18.16	May + June	25.99	16.49	+2.63	+1.67
	June	29.03	18.42	May + June	25.99	16.49	+3.04	+1.93
	June	29.03	18.42	June + July	26.17	16.61	+2.86	+1.81
	July	28.35	17.99	June + July	26.17	16.61	+2.18	+1.38
	July	28.35	17.99	July + Aug.	24.88	15.79	+3.47	+2.20
	August	26.87	17.05	July + Aug.	24.88	15.79	+1.99	+1.26
	August	26.87	17.05	Aug. + Sept.	23.18	14.71	+3.69	+2.34
	September	24.78	15.72	Aug. + Sept.	23.18	14.71	+1.60	+1.01
	September	24.78	15.72	Sept. + Oct.	23.93	15.18	+0.85	+0.54
	October	27.37	17.37	Sept. + Oct.	23.93	15.18	+3.44	+2.19

The average deviations without regard to sign are of course much larger since they constitute a measure of the error of prediction of the records of individual birds. They range from 24.8 to 30.1 eggs. The significance of errors of this magnitude will be more clearly brought out later.

The square root of mean square deviation also shows considerable regularity from month to month. These measures are naturally considerably larger than the average deviation without regard to sign. They range from 32.9 to 38.8 eggs.

It is clear that the annual egg production of birds similar in origin to the series upon which the prediction equations were based and maintained under similar conditions may be predicted with a relatively high degree of accuracy providing their record for any month is definitely known.

The accuracy with which prediction may be made will be clear if the errors of prediction are expressed in terms of the actual average annual production of the group of birds upon which the test is made.

TABLE 4

Square root of mean square deviation of predicted annual egg record from actual record. Prediction of annual production from one- and from two-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM TWO MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFER-ENCE IN PERCENT-AGE DEVIATION
	Base of prediction	Actual deviation	Percent-age deviation	Base of prediction	Actual deviation	Percent-age deviation		
For the whole year	November	38.65	24.52	Nov. + Dec.	36.46	23.13	+2.19	+1.39
	December	37.61	23.86	Nov. + Dec.	36.46	23.13	+1.15	+0.73
	December	37.61	23.86	Dec. + Jan.	35.50	22.53	+2.11	+1.33
	January	38.77	24.60	Dec. + Jan.	35.50	22.53	+3.27	+2.07
	January	38.77	24.60	Jan. + Feb.	34.32	21.78	+4.45	+2.82
	February	34.70	22.02	Jan. + Feb.	34.32	21.78	+0.38	+0.24
	February	34.70	22.02	Feb. + Mar.	31.32	19.87	+3.38	+2.15
	March	34.28	21.75	Feb. + Mar.	31.32	19.87	+2.96	+1.88
	March	34.28	21.75	Mar. + Apr.	32.89	20.87	+1.39	+0.88
	April	35.31	22.40	Mar. + Apr.	32.89	20.87	+2.42	+1.53
	April	35.31	22.40	Apr. + May	32.76	20.79	+2.55	+1.61
	May	35.89	22.77	Apr. + May	32.76	20.79	+3.13	+1.98
	May	35.89	22.77	May + June	32.53	20.64	+3.36	+2.13
	June	36.53	23.18	May + June	32.53	20.64	+4.00	+2.54
	June	36.53	23.18	June + July	33.00	20.94	+3.53	+2.24
	July	35.89	22.77	June + July	33.00	20.94	+2.89	+1.83
	July	35.89	22.77	July + Aug.	31.83	20.20	+4.06	+2.57
	August	34.34	21.79	July + Aug.	31.83	20.20	+2.51	+1.59
	August	34.34	21.79	Aug. + Sept.	30.39	19.28	+3.95	+2.51
	September	32.94	20.90	Aug. + Sept.	30.39	19.28	+2.55	+1.62
	September	32.94	20.90	Sept. + Oct.	32.74	20.77	+0.20	+0.13
	October	36.47	23.14	Sept. + Oct.	32.74	20.77	+3.73	+2.37

Remembering that the average annual production of the 415 test birds is 157.573 eggs, we use this as a base to determine the percentage errors for the equations for each month. These are given in columns with the caption "percentage deviation" in the tables.

We note that in predicting from December, February and August record the average error with regard to sign *is less than one percent of the average annual yield of the flock.* In predicting from November, January

nd March the error lies between one and two percent. When April, May, June, July, September and October records are used as a basis the verage errors of prediction are about 2.50 to 4.50 percent of the average nnual yield.

The average deviations without regard to sign are less than 20 percent f the annual production. The values for the individual months range rom 15.7 for September to 19.1 for January.

The square root of mean square deviations are less than 25 percent of he average annual production. The individual values range from 20.9 or September to 24.6 for January.

These two latter tests may at first seem to indicate very unsatisfactory rediction. Such, however, is not the case. These give the average rrors *either above or below the true record made in the prediction of the results or an individual bird.* The thing which is required in practice is generally *he prediction for a group of birds of a particular class.* In a flock of 415 irds this has been shown above to be possible *with an error of less than 5 ercent of the actual production for any month of the year and less than one ercent for a number of the months.*

The closeness of prediction may be made clear by a set of diagrams.

In these the estimated production is shown by the straight line. The ctual average production for the year or for the group of remaining months or which prediction is made is shown by solid dots for each group of birds s classified by monthly record. The shaded areas are determined as ollows. The birds were first grouped into classes of five-eggs range with espect to number of eggs laid during the period of time used as a basis f prediction. The birds of these classes of five-eggs range were further ıbdivided into those in which actual egg production was greater than the redicted and those in which the actual number was less than the predicted umber.[4] The average error of prediction was determined for each of ıese groups, and these averages represent the upper and the lower limits f the shaded areas. The upper limit represents, therefore, the average eviation (for the period for which prediction is made) of all birds which ıake a higher record than that predicted for their class. The lower limit

---

[4] A range of five eggs was used in order to obtain a number of birds sufficiently large to reduce mewhat the irregularities due to the errors of random sampling. The errors of prediction were each case determined for classes of unit range. Grouping is used for graphic representation erely. The average deviations represented by the limits of the shaded zone are to be thought as measured from a line perpendicular to the ordinates and intersecting the prediction line ı the mid-ordinate of the 5-egg class.

of the shaded area marks the average deviation for all birds which show an egg record lower than that predicted.

The graphs representing the prediction of annual production from the individual-months production appear in diagram 1 for the first six months of the year and in diagram 2 for the last half of the year.

DIAGRAM 1.—Tests of prediction of annual production from single-month records. Month of November to April. For explanation see text.

Notwithstanding the irregularities which are inevitable in graphs based on such a highly variable character as annual egg production in a flock of only 415 birds, the most critical reader must admit that the prediction is excellent.

Prediction of the production of a group of remaining months from the record of any month

As noted above (pages 266–268) the worker may desire to predict either the total egg production for the year or the egg production for a group of subsequent months of the year.

DIAGRAM 2.—Tests of prediction of annual production from single-month records. Tests for May to October.

In general the requirement will probably be the prediction of the total egg production of the remaining months of the year. Since, however, it is necessary to deal with other groups later, the errors of prediction of (a) the total egg production of the months of the year subsequent to the pth

month, where p is the base of prediction and of (b) the months of the year subsequent to the $(p + 1)$th month will be considered in this place.[5]

The equations required are as follows:

Month from which pre-diction is made	Period for which prediction is made	Prediction equation
November	December to October	$E_{11} = +143.186 + 1.914\,e_1$
November	January to October	$E_{10} = +139.262 + 1.403\,e_1$
December	January to October	$E_{10} = +134.491 + 1.835\,e_2$
December	February to October	$E_9 = +131.461 + 1.373\,e_2$
January	February to October	$E_9 = +130.997 + 1.564\,e_3$
January	March to October	$E_8 = +123.011 + 1.215\,e_3$
February	March to October	$E_8 = +109.824 + 2.035\,e_4$
February	April to October	$E_7 = +96.619 + 1.614\,e_4$
March	April to October	$E_7 = +69.966 + 2.471\,e_5$
March	May to October	$E_6 = +60.338 + 1.932\,e_5$
April	May to October	$E_6 = +46.490 + 2.523\,e_6$
April	June to October	$E_5 = +39.849 + 1.786\,e_6$
May	June to October	$E_5 = +27.639 + 2.233\,e_7$
May	July to October	$E_4 = +20.623 + 1.581\,e_7$
June	July to October	$E_4 = +13.895 + 1.920\,e_8$
June	August to October	$E_3 = +8.740 + 1.228\,e_8$
July	August to October	$E_3 = +6.049 + 1.440\,e_9$
July	September to October	$E_2 = +2.323 + 0.746\,e_9$
August	September to October	$E_2 = +0.724 + 0.935\,e_{10}$
August	October	$E_1 = +0.407 + 0.264\,e_{10}$
September	October	$E_1 = -0.726 + 0.480\,e_{11}$

The test of accuracy of prediction of these equations when applied to the 415 White Leghorns of 1917–1918 is given in comparison with the results for the prediction from two-months production (to be discussed later) in tables 5 to 7.

Limiting our attention for the moment to the errors of predicting the production of the months of the year remaining after any given month used as a basis of prediction, we note that in general the average deviations

---

[5] In the comparison between the egg production of a period of two months and the egg production of a single month as a basis of prediction, it is necessary to base critical comparisons upon the results of predictions of the records of periods subsequent to the two months under consideration. Concretely, if we are to compare November-plus-December record with November record and with December record as bases for the prediction of the annual production, the two-month period will contribute more to the annual record than either of the two months individually considered. Neither will contribute to the January-to-October production. We must, therefore, in testing prediction equations, base the test upon the results secured in predicting January-to-October egg record.

For this purpose we must have equations which show the relation between the egg record of the individual months and the egg record of groups of remaining months. For example, we require for November the January-to-October production; for December, the February-to-October production; for January, the March-to-October production and so on. For convenience merely the equations are given here in comparison with the other one-month equations.

TABLE 5

Average deviation with regard to sign of egg record for a period of months from actual record. Prediction from one- and from two-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM TWO MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
Dec. to Oct......	November	+2.39	1.57	—	—	—	—	—
Jan. to Oct......	November	+1.73	1.19	Nov. + Dec.	+1.16	0.80	+0.57	+0.39
Jan. to Oct......	December	+0.24	0.16	Nov. + Dec.	+1.16	0.80	−0.92	−0.64
Feb. to Oct......	December	+0.98	0.71	Dec. + Jan.	+1.69	1.22	−0.71	−0.51
Feb. to Oct......	January	+2.37	1.71	Dec. + Jan.	+1.69	1.22	+0.68	+0.49
Mar. to Oct......	January	+2.02	1.57	Jan. + Feb.	+1.45	1.12	+0.57	+0.45
Mar. to Oct......	February	+0.09	0.77	Jan. + Feb.	+1.45	1.12	−1.36	−0.35
Apr. to Oct......	February	+0.56	0.50	Feb. + Mar.	−0.05	0.04	+0.51	+0.46
Apr. to Oct......	March	−0.24	0.21	Feb. + Mar.	−0.05	0.04	+0.19	+0.17
May to Oct......	March	−1.51	1.60	Mar. + Apr.	−3.09	3.27	−1.58	−1.67
May to Oct......	April	−4.00	4.23	Mar. + Apr.	−3.09	3.27	+0.91	+0.96
June to Oct......	April	−1.70	2.33	Apr. + May	+0.64	0.88	+1.06	+1.45
June to Oct......	May	+3.62	4.98	Apr. + May	+0.64	0.88	+2.98	+4.10
July to Oct......	May	+1.55	2.90	May + June	−0.45	0.84	+1.10	+2.06
July to Oct......	June	−2.86	5.34	May + June	−0.45	0.84	+2.41	+4.50
Aug. to Oct......	June	−3.32	9.33	June + July	−3.73	10.48	−0.41	−1.15
Aug. to Oct......	July	−3.71	10.43	June + July	−3.73	10.48	−0.02	−0.05
Sept. to Oct......	July	−3.19	16.89	July + Aug.	−3.32	17.60	−0.13	−0.71
Sept. to Oct......	August	−2.56	13.57	July + Aug.	−3.32	17.60	−0.76	−4.03
October..........	August	−1.10	18.74	Aug. + Sept.	−0.79	13.46	+0.31	+5.28
October..........	September	−0.45	7.67	Aug. + Sept.	−0.79	13.46	−0.34	−5.79

TABLE 6

Average deviation without regard to sign of predicted egg record for a period of months from actual record. Prediction from one- and from two-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM TWO MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
Dec. to Oct.	November	29.59	19.49	—	—	—	—	—
Jan. to Oct.	November	28.47	19.56	Nov. + Dec.	28.09	19.30	+0.38	+0.26
Jan. to Oct.	December	28.43	19.53	Nov. + Dec.	28.09	19.30	+0.34	+0.23
Feb. to Oct.	December	26.92	19.38	Dec. + Jan.	26.25	18.90	+0.67	+0.48
Feb. to Oct.	January	26.48	19.06	Dec. + Jan.	26.25	18.90	+0.23	+0.16
Mar. to Oct.	January	25.03	19.40	Jan. + Feb.	24.98	19.37	+0.05	+0.03
Mar. to Oct.	February	24.07	18.66	Jan. + Feb.	24.98	19.37	−0.91	−0.71
Apr. to Oct.	February	22.93	20.46	Feb. + Mar.	22.39	19.98	+0.54	+0.48
Apr. to Oct.	March	22.81	20.36	Feb. + Mar.	22.39	19.98	+0.42	+0.38
May to Oct.	March	21.78	23.03	Mar. + Apr.	21.31	22.54	+0.47	+0.49
May to Oct.	April	21.36	22.59	Mar. + Apr.	21.31	22.54	+0.05	+0.05
June to Oct.	April	20.22	27.80	Apr. + May	19.65	27.02	+0.57	+0.78
June to Oct.	May	19.89	27.35	Apr. + May	19.65	27.02	+0.24	+0.33
July to Oct.	May	18.09	33.79	May + June	17.30	32.31	+0.79	+1.48
July to Oct.	June	17.62	32.91	May + June	17.30	32.31	+0.32	+0.60
Aug. to Oct.	June	14.88	41.82	June + July	14.05	39.49	+0.83	+2.33
Aug. to Oct.	July	13.91	39.09	June + July	14.05	39.49	−0.14	−0.40
Sept. to Oct.	July	10.92	57.90	July + Aug.	10.36	54.93	+0.56	+2.97
Sept. to Oct.	August	9.76	51.75	July + Aug.	10.36	54.93	−0.60	−3.18
October	August	5.67	96.59	Aug. + Sept.	5.13	87.39	+0.54	+9.20
October	September	4.57	77.85	Aug. + Sept.	5.13	87.39	−0.56	−9.54

TABLE 7

Square root of mean square deviation of predicted egg record from actual record. Prediction from one- and from two-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM TWO MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
Dec. to Oct........	November	38.65	25.46	—	—	—	—	—
Jan. to Oct........	November	37.05	25.45	Nov. + Dec.	36.46	25.05	+0.59	+0.40
Jan. to Oct........	December	36.63	25.16	Nov. + Dec.	36.46	25.05	+0.17	+0.11
Feb. to Oct........	December	34.94	25.15	Dec. + Jan.	34.15	24.58	+0.79	+0.57
Feb. to Oct........	January	34.61	24.91	Dec. + Jan.	34.15	24.58	+0.46	+0.33
Mar. to Oct........	January	32.52	25.22	Jan. + Feb.	32.28	25.03	+0.24	+0.19
Mar. to Oct........	February	30.86	23.92	Jan. + Feb.	32.28	25.03	−1.42	−1.11
Apr. to Oct........	February	29.17	26.03	Feb. + Mar.	28.12	25.09	+1.05	+0.94
Apr. to Oct........	March	28.40	25.34	Feb. + Mar.	28.12	25.09	+0.28	+0.25
May to Oct........	March	27.09	28.65	Mar. + Apr.	26.51	28.03	+0.58	+0.62
May to Oct........	April	26.50	28.02	Mar. + Apr.	26.51	28.03	−0.01	−0.01
June to Oct........	April	25.02	34.40	Apr. + May	24.48	33.66	+0.54	+0.74
June to Oct........	May	24.81	34.11	Apr. + May	24.48	33.66	+0.33	+0.45
July to Oct........	May	22.34	41.72	May + June	21.24	39.67	+1.10	+2.05
July to Oct........	June	21.44	40.04	May + June	21.24	39.67	+0.20	+0.37
Aug. to Oct........	June	18.18	51.09	June + July	17.40	48.90	+0.78	+2.19
Aug. to Oct........	July	17.14	48.17	June + July	17.40	48.90	−0.26	−0.73
Sept. to Oct........	July	13.21	70.02	July + Aug.	12.67	67.18	+0.54	+2.84
Sept. to Oct........	August	12.12	64.26	July + Aug.	12.67	67.18	−0.55	−2.92
October............	August	6.92	117.84	Aug. + Sept.	6.24	106.30	+0.68	+11.54
October............	September	5.71	97.27	Aug. + Sept.	6.24	106.30	−0.53	−9.03

with regard to sign are small. No one of the errors is over 4 eggs. Th
percentage values, in which the actual average yields of the remainin
months in question are used as bases, range from 0.2 for the prediction c
January-to-October production from December production to 13.6 percen
in the case of the prediction of September-to-October production from th
August record. The average deviations without regard to sign range fror
4.6 to 29.6 eggs. The percentage values range from 18.7 to 77.9 percen
of the actual production for the given remaining period.

DIAGRAM 3.—Tests of prediction of production for a group of remaining months from single
month records. Tests for November to April. For explanation see text.

The square root of mean square deviations vary from 5.7 to 38.7 eggs
or from 23.9 to 97.3 percent of the actual yield.

The values of the average deviation without regard to sign and o
square root of mean square deviation decrease from the earlier to the late
months. This is, of course, due to the fact that in predicting the eg
record of the remaining months of the year the total record decreases a
the number of remaining months becomes smaller. It is to be expecte

DIAGRAM 4.—Tests of prediction of the production of a group of remaining months from single-month production.   Tests for May to September.

therefore, that the absolute error of prediction will be smaller than when the prediction is made for a longer period. The relative errors of prediction are conspicuously larger than those found when the prediction is made for the year as a whole. Furthermore, these relative (percentage) errors increase as the period for which prediction is made becomes shorter. The test shows clearly that prediction of the results of short remaining periods cannot be made,—at least by means of the linear equations for prediction from one month's record tested in this paper,—with a satisfactory degree of accuracy.

When prediction is made for the period subsequent to the $(p + 1)$th month the average deviations with regard to sign vary from 0.56 to 3.32 or from 0.50 to 18.74 percent of the actual production for the period. The average deviations without regard to sign vary from 28.47 eggs for the prediction of January-to-October production from November production to 5.67 eggs for the prediction of October production from August record. The percentage values range from 19.38 to 96.59 percent. Similar results are found in the case of the square root of mean square deviation which ranges from 37.05 eggs for the prediction of January-to-October production from November record to 6.92 eggs for the prediction of October production from August record. The percentage values range from 25.2 to 117.8 percent of the actual records.

The graphic representation of the errors of the prediction of the remaining months of the year is made in diagrams 3 and 4.

The slope of the lines and the moderate narrowness of the shaded areas as well as the fair agreement of the empirical and the predicted means for the remaining periods, evidence for fairly satisfactory prediction for the first six months of the year. As the end of the year is approached and as the period of remaining months becomes shorter the slopes of the lines are more moderate. The narrowness of the shaded areas, representing the difference between the averages of the errors of over-prediction and under-prediction, does not indicate great accuracy of prediction as compared with that attainable in the earlier months, but merely that (because of the smaller egg record made by birds in the latter months of the year) great deviations from prediction are improbable. It is evident, therefore, that for the prediction of the record of the later months of the year from the record of immediately preceding months the equations have relatively little value.

It is quite clear that while the prediction of a group of remaining months may be made with a relatively high degree of accuracy early in the year the predictions are relatively poor toward the end of the year.

Prediction of annual production from the sum of two monthly records

Before considering the results of equations for the prediction of annual production from the combined record of two or more months, some general questions of theory must be considered.

If the egg production of each individual month be correlated with that of the whole year it would seem that a better prediction of the annual total may be made from the record of two or more individual monthly records than from one month's record only. This is a point emphasized by CARD (1917) who has correlated the total production of groups of months with the annual yield.

There are several points to be taken into consideration here. First, it should be clear that the superiority of a group of months for predicting the annual yield of a bird is to a considerable extent due to the fact that the records of these months are included in the annual total. Thus in predicting annual total from November performance, the November record is included in the annual total. In predicting from November, December and January production the records of these *three* months are included in the annual total. As far as their own contribution is concerned, prediction can be made with absolute certainty. The importance of this factor would be especially great during the spring months when the number of eggs laid by practically all birds is high. If the principle of an increase in the number of months upon which prediction is to be based be extended to its limit, it is clear that the annual total can be predicted with exactness from the record of twelve months. The importance of this factor was fully recognized in our second publication (HARRIS, BLAKESLEE and KIRK-PATRICK 1918), in which we determined the correlation between the production of each individual month and that of the remaining eleven months of the year, as well as that between the production of the individual months and the annual record.

It is evident that it is impossible to compare directly and critically the errors made in predicting annual egg production from two-month periods and from single-month periods; in one case a single component only is included in the first and second variable of the pair whereas in the second case two components are involved. The problem of a direct comparison will be taken up in a subsequent section.

Second, from the economic standpoint it is clear that trap-nesting for two months or three months is (disregarding initial investment) twice or three times as expensive as trap-nesting for one month. In general it is important to utilize the shortest practicable period on which prediction may be based.

Third, the mathematical theory of multiple correlation shows that in dealing with correlated characters the gain in accuracy of prediction rapidly decreases with the number of characters employed. In our first detailed treatment of the problem of the correlation between the egg records of the individual months we showed by the constants for a series of selected months that the egg records of the individual months are correlated among themselves. This has since been demonstrated for the entire series of $\frac{1}{2} n(n-1) = 66$ different combinations of the 12 months of the pullet year. It is evident, therefore, that very large gains in accuracy of prediction cannot be expected to result from an increase in the number of periods, except in so far as the gain is due directly to the contribution of the months included.

We now turn to the results of the test of equations for the prediction of annual record from two-consecutive-months production. The equations are as follows:

Months from which prediction is made	Prediction equation
November and December	$E = +132.887 + 2.160\,(e_1 + e_2)$
December and January	$E = +130.822 + 2.176\,(e_2 + e_3)$
January and February	$E = +146.040 + 2.579\,(e_3 + e_4)$
February and March	$E = +78.008 + 2.915\,(e_4 + e_5)$
March and April	$E = +48.374 + 3.029\,(e_5 + e_6)$
April and May	$E = +39.955 + 3.005\,(e_6 + e_7)$
May and June	$E = +32.783 + 3.065\,(e_7 + e_8)$
June and July	$E = +44.650 + 2.864\,(e_8 + e_9)$
July and August	$E = +62.861 + 2.625\,(e_9 + e_{10})$
August and September	$E = +91.865 + 2.302\,(e_{10} + e_{11})$
September and October	$E = +122.597 + 2.140\,(e_{11} + e_{12})$

Since a primary object of the present analyses is a comparison of equations based on two-months production with those based on a single month's record as a means of predicting the annual egg record of a bird, it is advantageous to place the results for the two methods side by side in the same tables. The results are given in tables 2 to 4.[6]

Table 2 shows the average errors with regard to sign of the egg records of the 415 White Leghorns studied at Storrs in 1917–1918, when prediction is made from two-months production using equations based on the Storrs experience of the preceding six years.

The average deviations with regard to sign are small. In 7 cases the equations have predicted values which are too large, whereas in 4 cases they have predicted values which are too small. The individual errors

---

[6] It has seemed conducive to clearness to duplicate entries in order to secure the 22 differences which serve as a basis of comparison.

e very small. Two are less than 1 egg, 4 are less than 2 eggs, while 5
e from 2 to 7 eggs. The percentage errors based on the mean annual
oduction are less than 1 percent in 5 of the cases and less than 5 percent

DIAGRAM 5.—Tests of prediction of annual production from combined record of two consecu-
te months. For explanation see text. Tests for November to February.

i the other 6 cases. The average error in actual number of eggs, dis-
ıgarding the sign of the error, is 2.72 eggs while the average of percentage
ırors is 1.72 percent.

It seems unnecessary to discuss in detail the average deviations without
·gard to sign, of the predicted from the observed annual egg production.

The errors, shown in table 3, range from 23.2 to 28.1 eggs or from 14 to 17.8 percent.

Similar results for the square root of mean square deviation are given in table 4 which shows that prediction from the sum of two-consecutive

DIAGRAM 6.—Tests of prediction of annual production from the combined record of two consecutive months.   Tests of February-to-May production.

months production gives a square root of mean square deviation ranging from 30.4 to 36.5 eggs or from 19.3 to 23.1 percent of the annual production.

Thus it is clear that the annual egg record of a bird may be predicted with a high degree of accuracy from the combined egg record of any two consecutive months.

These results may be represented graphically by diagrams 5–8, which
ve been prepared on the same principle as those for the results of predic-
on from a single month's production.   The general excellence of the agree-
nt (considering the fact that there are only 415 birds upon which equa-

DIAGRAM 7.—Tests of the prediction of annual production from the combined record of two
consecutive months.   Tests for May to August.

ins based upon an entirely different series are being tested) renders
tailed discussion of these diagrams superfluous.

The most interesting feature of these tables is, however, the comparison
tween the value of two-months observation and of single-month obser-
vtion as bases for the estimation of the total (annual) egg-producing
cpacity of the organism.

The differences in the average deviations with regard to sign, as show in the first of the two final columns of table 2, range from less than a sing egg (5 comparisons) to a maximum of less than 7 eggs. The avera difference is 2.21 eggs. If signs be considered the average difference only + 0.67 eggs. The differences in the percentage deviation when predi tion is made by single- and by two-month periods are shown in the fin column of the table.

DIAGRAM 8.—Tests of prediction of annual production from the combined record of two co secutive months. Tests of August to October.

It is clear from these results that the results of prediction from tw months production are not materially better from the practical stan point than those for single-month's production although the labor entail in recording the performance of a bird for two months must be appro> mately twice as great as that for a single month.

The reader who cares to do so may verify these statements by a stud of the results for average deviation without regard to sign and for squa root of mean square deviation as shown in the two final columns of tabl 3 and 4.

Prediction of the production of a group of remaining months from the sum of two monthly records

We now have to consider the problem of the accuracy with which the egg production of a group of subsequent months may be predicted from the sum of two consecutive monthly records.

The equations are the following:

Period from which prediction is made	Period for which prediction is made	Prediction equation
November and December	January to October	$E = +132.887 + 1.160\ (e_1 + e_2)$
December and January	February to October	$E = +128.112 + 0.979\ (e_2 + e_3)$
January and February	March to October	$E = +124.959 + 0.336\ (e_3 + e_4)$
February and March	April to October	$E = +76.542 + 1.320\ (e_4 + e_5)$
March and April	May to October	$E = +44.586 + 1.363\ (e_5 + e_6)$
April and May	June to October	$E = +25.022 + 1.231\ (e_6 + e_7)$
May and June	July to October	$E = +7.280 + 1.118\ (e_7 + e_8)$
June and July	August to October	$E = +0.994 + 0.827\ (e_8 + e_9)$
July and August	September to October	$E = -7.281 + 0.660\ (e_9 + e_{10})$
August and September	October	$E = -2.137 + 0.245\ (e_{10} + e_{11})$

The results appear in the second section of tables 5 to 7. Here they are laid beside the errors obtained for the prediction of the production of these same periods from the record of the two months individually considered, as given by the equations shown on page 282.

Table 5, giving the average deviation with regard to sign of the predicted from the observed values, shows that the actual deviations have a numerical range of 0.05 to 3.73 eggs or from 0.04 to 17.6 percent. The largest relative (percentage) deviations are, of course, in the final months of the year.

The average deviations without regard to sign appear in the second column of table 6. These vary from as low as 5.13 eggs in October to 28.09 eggs for the period January to October. Since the average production decreases as the number of remaining months becomes smaller we find the largest percentage errors in the later groups of months. These percentage values range from 18.9 for the period February to October to 87.4 for the month of October. Similar results with somewhat different numerical values are found in table 7 which shows the square root of mean square deviation of the predicted from the observed values.

These results show that when the number of remaining months is large, prediction of egg production can be made with relatively high accuracy from the combined record of two months. As the number of months becomes smaller the error of prediction is, as compared with the average production, relatively large.

Turning now to the problem of the comparison of periods of one month and of two months as bases of prediction, and testing the efficiency of these two periods on the egg production of comparable remaining periods of time, we note that the differences in the two final columns of tables 5 to 7, expressed either in number of eggs or in percentages of the total production, are small. Thus the differences for the average deviation with regard to sign are all less than 3 eggs and all less than 6 percent. Most of the differences are far smaller than this. In some cases the prediction from a single month gives the better result; in others prediction from two months gives the better result. The differences in the errors without regard to sign as obtained by the two methods are even smaller. No difference amounts to as much as a single egg per year. The large differences in the percentage errors by the two methods are found exclusively in the later months of the year where the total production is low. Comparable, but numerically somewhat different, results are found for the square root of mean square deviation.

Thus it is clear that there is little practical difference between single-month and two-months production as bases of the prediction of the egg record of a subsequent period.

Prediction of annual production from the sum of three monthly records

The equations required for the prediction of annual production from the combined record of three consecutive months are the following:

Months from which prediction is made	Prediction equation
November, December and January	$E = +126.742 + 1.770 \, (e_1 + e_2 + e_3)$
December, January and February	$E = +113.940 + 1.951 \, (e_2 + e_3 + e_4)$
January, February and March	$E = +82.129 + 2.266 \, (e_3 + e_4 + e_5)$
February, March and April	$E = +50.502 + 2.323 \, (e_4 + e_5 + e_6)$
March, April and May	$E = +29.450 + 2.267 \, (e_5 + e_6 + e_7)$
April, May and June	$E = +19.349 + 2.324 \, (e_6 + e_7 + e_8)$
May, June and July	$E = +23.786 + 2.233 \, (e_7 + e_8 + e_9)$
June, July and August	$E = +41.079 + 2.065 \, (e_8 + e_9 + e_{10})$
July, August and September	$E = +67.078 + 1.895 \, (e_9 + e_{10} + e_{11})$
August, September and October	$E = +97.699 + 1.794 \, (e_{10} + e_{11} + e_{12})$

The second section of table 8 shows the average deviation with regard to sign of the annual egg production predicted from the combined record of 3 consecutive months from the performance of the 415 White Leghorn birds studied at Storrs in 1917–1918.

The results show that the trimonthly totals, like the monthly records and bimonthly totals considered in preceding sections, give excellent predictions. December to February, January to March, March to May, and

uly to September give average errors of prediction of less than 1 egg. November to January, April to June, and May to July give errors of prediction of between 2 and 3 eggs. August to October gives an error of predic-

TABLE 8

verage deviation with regard to sign of predicted annual egg record from actual record. Prediction of annual production from one- and from three-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM THREE MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
or the whole year	November	+2.39	1.52	Nov.–Jan.	+2.09	1.33	+0.30	+0.19
	December	−0.49	0.31	Nov.–Jan.	+2.09	1.33	−1.60	−1.02
	January	+2.58	1.64	Nov.–Jan.	+2.09	1.33	+0.49	+0.31
	December	−0.49	0.31	Dec.–Feb.	+0.78	0.49	−0.29	−0.18
	January	+2.58	1.64	Dec.–Feb.	+0.78	0.49	+1.80	+1.15
	February	+0.06	0.04	Dec.–Feb.	+0.78	0.49	−0.72	−0.45
	January	+2.58	1.64	Jan.–Mar.	+0.49	0.31	+2.09	+1.33
	February	+0.06	0.04	Jan.–Mar.	+0.49	0.31	−0.43	−0.27
	March	−1.63	1.03	Jan.–Mar.	+0.49	0.31	+1.14	+0.72
	February	+0.06	0.04	Feb.–Apr.	−4.07	2.58	−4.01	−2.54
	March	−1.63	1.03	Feb.–Apr.	−4.07	2.58	−2.44	−1.55
	April	−6.23	3.95	Feb.–Apr.	−4.07	2.58	+2.16	+1.37
	March	−1.63	1.03	Mar.–May	−0.73	0.46	+0.90	+0.57
	April	−6.23	3.95	Mar.–May	−0.73	0.46	+5.50	+3.49
	May	+7.02	4.45	Mar.–May	−0.73	0.46	+6.29	+3.99
	April	−6.23	3.95	Apr.–June	−2.31	1.47	+3.92	+2.48
	May	+7.02	4.45	Apr.–June	−2.31	1.47	+4.71	+2.98
	June	−5.21	3.31	Apr.–June	−2.31	1.47	+2.90	+1.84
	May	+7.02	4.45	May–July	−2.12	1.35	+4.90	+3.10
	June	−5.21	3.31	May–July	−2.12	1.35	+3.09	+1.96
	July	−5.27	3.34	May–July	−2.12	1.35	+3.15	+1.99
	June	−5.21	3.31	June–Aug.	−5.35	3.39	−0.14	−0.08
	July	−5.27	3.34	June–Aug.	−5.35	3.39	−0.08	−0.05
	August	−0.82	0.52	June–Aug.	−5.35	3.39	−4.53	−2.87
	July	−5.27	3.34	July–Sept.	−0.20	0.13	+5.07	+3.21
	August	−0.82	0.52	July–Sept.	−0.20	0.13	+0.62	+0.39
	September	+4.78	3.03	July–Sept.	−0.20	0.13	+4.58	+2.90
	August	−0.82	0.52	Aug.–Oct.	+3.91	2.48	−3.09	−1.96
	September	+4.78	3.03	Aug.–Oct.	+3.91	2.48	+0.87	+0.55
	October	+3.95	2.51	Aug.–Oct.	+3.91	2.48	+0.04	+0.03

on of between 3 and 4 eggs. February to April gives an error of prediction between 4 and 5 eggs. Finally June to August gives an error of prediction of between 5 and 6 eggs.

Considered in their relation to the average annual production these values range from 0.13 to 3.39 percent. These results certainly show remarkable accuracy of prediction.

The average errors without regard to sign, given in table 9 need not be considered in detail.   They range from 21.4 to 25.9 eggs per year or from 13.6 to 16.5 percent of the annual total.

TABLE 9

Average deviation without regard to sign of predicted annual egg record from actual record. Prediction of annual production from one- and from three-months performance.   Equations based on Storrs experience, 1911 to 1917.   Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM THREE MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
	November	29.59	18.78	Nov.–Jan.	25.93	16.45	+3.66	+2.33
	December	29.26	18.57	Nov.–Jan.	25.93	16.45	+3.33	+2.12
	January	30.09	19.09	Nov.–Jan.	25.93	16.45	+4.16	+2.64
	December	29.26	18.57	Dec.–Feb.	25.31	16.06	+3.95	+2.51
	January	30.09	19.09	Dec.–Feb.	25.31	16.06	+4.78	+3.03
	February	27.28	17.31	Dec.–Feb.	25.31	16.06	+1.97	+1.25
	January	30.09	19.09	Jan.–Mar.	25.29	16.05	+4.80	+3.04
	February	27.28	17.31	Jan.–Mar.	25.29	16.05	+1.99	+1.26
	March	27.95	17.73	Jan.–Mar.	25.29	16.05	+2.66	+1.68
	February	27.28	17.31	Feb.–Apr.	24.16	15.33	+3.12	+1.98
	March	27.95	17.73	Feb.–Apr.	24.16	15.33	+3.79	+2.40
	April	28.72	18.22	Feb.–Apr.	24.16	15.33	+4.56	+2.89
	March	27.95	17.73	Mar.–May	25.42	16.13	+2.53	+1.60
For the whole year	April	28.72	18.22	Mar.–May	25.42	16.13	+3.30	+2.09
	May	28.62	18.16	Mar.–May	25.42	16.13	+3.20	+2.03
	April	28.72	18.22	Apr.–June	24.33	15.44	+4.39	+2.78
	May	28.62	18.16	Apr.–June	24.33	15.44	+4.29	+2.72
	June	29.03	18.42	Apr.–June	24.33	15.44	+4.70	+2.98
	May	28.62	18.16	May–July	24.20	15.36	+4.42	+2.80
	June	29.03	18.42	May–July	24.20	15.36	+4.83	+3.06
	July	28.35	17.99	May–July	24.20	15.36	+4.15	+2.63
	June	29.03	18.42	June–Aug.	23.49	14.90	+5.54	+3.52
	July	28.35	17.99	June–Aug.	23.49	14.90	+4.86	+3.09
	August	26.87	17.05	June–Aug.	23.49	14.90	+3.38	+2.15
	July	28.35	17.99	July–Sept.	21.36	13.55	+6.99	+4.44
	August	26.87	17.05	July–Sept.	21.36	13.55	+5.51	+3.50
	September	24.78	15.72	July–Sept.	21.36	13.55	+3.42	+2.17
	August	26.87	17.05	Aug.–Oct.	21.59	13.70	+5.28	+3.35
	September	24.78	15.72	Aug.–Oct.	21.59	13.70	+3.19	+2.02
	October	27.37	17.37	Aug.–Oct.	21.59	13.70	+5.78	+3.67

The square root of mean square deviation of errors of prediction given in table 10 are, of course, larger than the average deviations without regard to sign.   They vary from 28.1 to 33.8 eggs or from 17.8 to 21.5 percent of the annual production.

The range of variation in the egg production of three-month periods is
) wide that it is impossible because of the limitations of space to represent
ιe errors of prediction from three-month periods graphically for each of
ιe equations.

TABLE 10

*ιuare root of mean square deviation of predicted annual egg record from actual record.   Prediction
of annual production from one- and from three-months performance.   Equations based on Storrs
experience, 1911 to 1917.   Test of equations on 415 White Leghorns, Storrs, 1917–1918.*

ERIOD FOR WHICH REDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM THREE MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFER- ENCE IN PERCENT- AGE DEVIATION
	Base of prediction	Actual deviation	Percent- age deviation	Base of prediction	Actual deviation	Percent- age deviation		
	November	38.65	24.52	Nov.–Jan.	33.84	21.47	+4.81	+3.05
	December	37.61	23.86	Nov.–Jan.	33.84	21.47	+3.77	+2.39
	January	38.77	24.60	Nov.–Jan.	33.84	21.47	+4.93	+3.13
	December	37:61	23.86	Dec.–Feb.	32.65	20.72	+4.96	+3.14
	January	38.77	24.60	Dec.–Feb.	32.65	20.72	+6.12	+3.88
	February	34.70	22.02	Dec.–Feb.	32.65	20.72	+2.05	+1.30
	January	38.77	24.60	Jan.–Mar.	31.58	20.04	+7.19	+4.56
	February	34.70	22.02	Jan.–Mar.	31.58	20.04	+3.12	+1.98
	March	34.28	21.75	Jan.–Mar.	31.58	20.04	+2.70	+1.71
	February	34.70	22.02	Feb.–Apr.	29.77	18.89	+4.93	+3.13
	March	34.28	21.75	Feb.–Apr.	29.77	18.89	+4.51	+2.86
	April	35.31	22.40	Feb.–Apr.	29.77	18.89	+5.54	+3.51
	March	34.28	21.75	Mar.–May	31.14	19.76	+3.14	+1.99
ɔr the whole year	April	35.31	22.40	Mar.–May	31.14	19.76	+4.17	+2.64
	May	35.89	22.77	Mar.–May	31.14	19.76	+4.75	+3.01
	April	35.31	22.40	Apr.–June	30.59	19.41	+4.72	+2.99
	May	35.89	22.77	Apr.–June	30.59	19.41	+5.30	+3.36
	June	36.53	23.18	Apr.–June	30.59	19.41	+5.94	+3.77
	May	35.89	22.77	May–July	29.40	18.65	+6.49	+4.12
	June	36.53	23.18	May–July	29.40	18.65	+7.13	+4.53
	July	35.89	22.77	May–July	29.40	18.65	+6.49	+4.12
	June	36.53	23.18	June–Aug.	29.80	18.91	+6.73	+4.27
	July	35.89	22.77	June–Aug.	29.80	18.91	+6.09	+3.86
	August	34.34	21.79	June–Aug.	29.80	18.91	+4.54	+2.88
	July	35.89	22.77	July–Sept.	28.10	17.83	+7.79	+4.94
	August	34.34	21.79	July–Sept.	28.10	17.83	+6.24	+3.96
	September	32.94	20.90	July Sept.	28.10	17.83	+4.84	+3.07
	August	34.34	21.79	Aug.–Oct.	29.23	18.55	+5.11	+3.24
	September	32.94	20.90	Aug.–Oct.	29.23	18.55	+3.71	+2.35
	October	36.47	23.14	Aug.–Oct.	29.23	18.55	+7.24	+4.59

Two series, that for November to January and for March to May, have
ʰen selected at random to represent the goodness of fit of prediction in
ιese cases.   The results for prediction from November to January record
ιe shown in diagram 9.   Those for prediction from March to May pro-

DIAGRAM 9.—Tests of the prediction of annual production (upper figure) and of the production of a group of remaining months (lower figure) from the combined record of three consecutive months. Tests for the period November to January.

DIAGRAM 10.—Tests of the prediction of annual production (upper figure) and of the production of a group of remaining months (lower figure) from the combined record of three consecutive months. Tests for the period March to May.

PREDICTING EGG PRODUCTION IN WHITE LEGHORNS 301

uction are shown in diagram 10. In both cases the upper figure represents he prediction of annual production. The lower figure shows the prediction f the groups of remaining months and will be discussed in a subsequent ection.

After the discussion of the preceding diagrams these graphs are self-xplanatory.

When these results are compared, as in the last two columns of the tables, ith those for prediction from a single one of the three months the differ-nces are surprisingly small. For example the most important test,—that f the average deviation with regard to sign,—shows that 11 of the 30 ifferences are less than 1 egg per year; 3 are less than 2 eggs per year; hile 16 are 2 eggs or more per year. In no case is the difference as much s 7 eggs per year. The difference in percentage deviation is in no case s large as 4 percent.

Turning to the comparison of average deviation without regard to sign hen prediction is made from trimonthly periods and from the records f individual months we note that the differences are without exception ositive in sign. Thus they show a greater error when prediction is made om a single monthly record. The differences are, however, always less han 7 eggs per year and are generally less than 5 eggs. The percentage ifferences vary from 1.3 to 4.4 percent when both percentages are based n the annual total.

Similar results are obtained for the square root of mean square deviation. he deviations are larger throughout when prediction is made from single-onths records than when made from three-months records. The differ-nces are not, however, large. They range from 2.05 to 7.79 eggs, or from .30 to 4.94 percent of the annual average production.

Thus while practically without exception a closer prediction of the nnual egg record of individual birds can be made from three-months roduction the difference between a three-month period and a single-onth period is by no means so large as one unacquainted with statistical heory might have assumed.

ENETICS 6: My 1921

Prediction of the production of a subsequent period from the sum of three monthly records

The equations required are the following:

Months from which prediction is made	Period for which prediction is made	Prediction equation
November, Dec. and Jan.	February to October	$E_9 = +126.742 + 0.770\,(e_1 + e_2 + e_3)$
Dec., Jan. and Feb.	March to October	$E_8 = +112.051 + 0.806\,(e_2 + e_3 + e_4)$
Jan., Feb. and March	April to October	$E_7 = +81.464 + 0.935\,(e_3 + e_4 + e_5)$
February, March and April	May to October	$E_6 = +49.753 + 0.955\,(e_4 + e_5 + e_6)$
March, April and May	June to October	$E_5 = +25.210 + 0.850\,(e_5 + e_6 + e_7)$
April, May and June	July to October	$E_4 = +7.063 + 0.770\,(e_6 + e_7 + e_8)$
May, June and July	August to October	$E_3 = -1.975 + 0.594\,(e_7 + e_8 + e_9)$
June, July and August	September to October	$E_2 = -4.701 + 0.377\,(e_8 + e_9 + e_{10})$
July, August and September	October	$E_1 = -3.343 + 0.172\,(e_9 + e_{10} + e_{11})$

Table 11 contains the average deviations with regard to sign, of the predicted yield of remaining months, from the actual productions, when prediction is made from the total yield of three consecutive months.

The deviations range from 0.20 to 3.30 eggs or from 0.27 to 18.22 percent of the actually observed yield. As far as this criterion shows, predictions are excellent for all periods from that including February to October to that for August to October. The September-to-October record and the October record, however, cannot be predicted with a high degree of accuracy, the errors being over 17 percent of the mean value for these months.

The average deviations without regard to sign, shown in table 12, range from 5.24 to 25.93 eggs, the values decreasing as the length of the period for which prediction is made becomes smaller. The reverse is true of the percentage values which increase from 18.66 percent for the period February to October to 89.23 percent of the actual yield for the month of October.

Similar results are obtained when the formulae are judged by the square root of mean square deviation of the predicted from the actually observed egg record as shown in table 13. These root mean square deviations range from 33.84 for February to October to 6.44 for the month of October alone, or from 24.30 percent for the group of 8 remaining months of the year to 109.67 percent for the last (single) month.

The results for the prediction of two of the groups of remaining months from the combined records of three-months production are represented graphically for the three months November to January in diagram 9 and for the three months March to May in diagram 10. It is the lower figure which is to be consulted in each case.

The gentle slope of the lines and the considerable irregularities of the means show that prediction of the record of a period of remaining months

TABLE 11

Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM THREE MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
Feb. to Oct.........	November	+2.16	1.55	Nov.–Jan.	+2.09	1.50	+0.07	+0.05
Feb. to Oct.........	December	+0.99	0.71	Nov.–Jan.	+2.09	1.50	−1.10	−0.79
Feb. to Oct.........	January	+2.37	1.71	Nov.–Jan.	+2.09	1.50	+0.28	+0.21
Mar. to Oct.........	December	+0.91	0.71	Dec.–Feb.	+1.35	1.05	−0.44	−0.34
Mar. to Oct.........	January	+2.02	1.57	Dec.–Feb.	+1.35	1.05	+0.67	+0.52
Mar. to Oct.........	February	+0.99	0.77	Dec.–Feb.	+1.35	1.05	−0.36	−0.28
Apr. to Oct.........	January	+1.40	1.25	Jan.–Mar.	+0.73	0.65	+0.67	+0.60
Apr. to Oct.........	February	+0.56	0.50	Jan.–Mar.	+0.73	0.65	−0.17	−0.15
Apr. to Oct.........	March	−0.24	0.21	Jan.–Mar.	+0.73	0.65	−0.49	−0.44
May to Oct.........	February	−0.77	0.81	Feb.–Apr.	−2.51	2.65	−1.74	−1.84
May to Oct.........	March	−1.51	1.60	Feb.–Apr.	−2.51	2.65	−1.00	−1.05
May to Oct.........	April	−4.00	4.23	Feb.–Apr.	−2.51	2 65	+1.49	+1.58
June to Oct.........	March	−0.01	0.01	Mar.–May	+0.20	0.27	−0.19	−0.26
June to Oct.........	April	−1.70	2.33	Mar.–May	+0.20	0.27	+1.50	+2.06
June to Oct.........	May	+3.62	4.98	Mar.–May	+0.20	0.27	+3.42	+4.71
July to Oct.........	April	−2.16	4.03	Apr.–June	−1.49	2.78	+0.67	+1.25
July to Oct.........	May	+1.55	2.90	Apr.–June	−1.49	2.78	+0.06	+0.12
July to Oct.........	June	−2.86	5.34	Apr.–June	−1.49	2.78	+1.37	+2.56
Aug. to Oct.........	May	−0.51	1.43	May–July	−2.56	7.20	−2.05	−5.77
Aug. to Oct.........	June	−3.32	9.33	May–July	−2.56	7.20	+0.76	+2.13
Aug. to Oct.........	July	−3.71	10.43	May–July	−2.56	7.20	+1.15	+3.23
Sept. to Oct.........	June	−3.00	15.91	June–Aug.	−3.30	17.50	−0.30	−1.59
Sept. to Oct.........	July	−3.19	16.89	June–Aug.	−3.30	17.50	−0.11	−0.61
Sept. to Oct.........	August	−2.56	13.57	June–Aug.	−3.30	17.50	−0.74	−3.93
October............	July	−1.33	22.64	July–Sept.	−1.07	18.22	+0.26	+4.42
October............	August	−1.10	18.74	July–Sept.	−1.07	18.22	+0.03	+0.52
October............	September	−0.45	7.67	July–Sept.	−1.07	18.22	−0.62	−10.55

performance.

TABLE 12

Average deviation without regard to sign of predicted egg record for a period of months from actual record. Prediction from one- and from three-months performance. Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM THREE MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
Feb. to Oct.	November	26.67	19.20	Nov.–Jan.	25.93	18.66	+0.74	+0.54
Feb. to Oct.	December	26.92	19.38	Nov.–Jan.	25.93	18.66	+0.99	+0.72
Feb. to Oct.	January	26.48	19.06	Nov.–Jan.	25.93	18.66	+0.55	+0.40
Mar. to Oct.	December	25.33	19.61	Dec.–Feb.	24.19	18.75	+1.14	+0.86
Mar. to Oct.	January	25.03	19.40	Dec.–Feb.	24.19	18.75	+0.84	+0.65
Mar. to Oct.	February	24.07	18.66	Dec.–Feb.	24.19	18.75	−0.12	−0.09
Apr. to Oct.	January	23.68	21.13	Jan.–Mar.	22.63	20.20	+1.05	+0.93
Apr. to Oct.	February	22.93	20.46	Jan.–Mar.	22.63	20.20	+0.30	+0.26
Apr. to Oct.	March	22.81	20.36	Jan.–Mar.	22.63	20.20	+0.18	+0.16
May to Oct.	February	21.73	22.98	Feb.–Apr.	21.03	22.24	+0.70	+0.74
May to Oct.	March	21.78	23.03	Feb.–Apr.	21.03	22.24	+0.75	+0.79
May to Oct.	April	21.36	22.59	Feb.–Apr.	21.03	22.24	+0.33	+0.35
June to Oct.	March	20.57	28.28	Mar.–May	19.75	27.16	+0.82	+1.12
June to Oct.	April	20.22	27.80	Mar.–May	19.75	27.16	+0.47	+0.64
June to Oct.	May	19.89	27.35	Mar.–May	19.75	27.16	+0.14	+0.19
July to Oct.	April	18.49	34.54	Apr.–June	17.44	32.58	+1.05	+1.96
July to Oct.	May	18.09	33.79	Apr.–June	17.44	32.58	+0.65	+1.21
July to Oct.	June	17.62	32.91	Apr.–June	17.44	32.58	+0.18	+0.33
Aug. to Oct.	May	15.03	42.25	May–July	14.03	39.44	+1.00	+2.81
Aug. to Oct.	June	14.88	41.83	May–July	14.03	39.44	+0.85	+2.39
Aug. to Oct.	July	13.91	39.09	May–July	14.03	39.44	−0.12	−0.35
Sept. to Oct.	June	11.27	59.76	June–Aug.	10.26	54.40	+1.01	+5.36
Sept. to Oct.	July	10.92	57.90	June–Aug.	10.26	54.40	+0.66	+3.50
Sept. to Oct.	August	9.76	51.75	June–Aug.	10.26	54.40	−0.50	−2.65
October	July	5.76	98.09	July–Sept.	5.24	89.23	+0.52	+8.86
October	August	5.67	96.59	July–Sept.	5.24	89.23	+0.43	+7.36
October	September	4.57	77.85	July–Sept.	5.24	89.23	−0.67	−11.38

TABLE 13

performance. *Equations based on Storrs experience, 1911 to 1917. Test of equations on 415 White Leghorns, Storrs, 1917–1918.*

PERIOD FOR WHICH PREDICTION IS MADE	PREDICTION FROM ONE MONTH			PREDICTION FROM THREE MONTHS			DIFFERENCE IN ACTUAL DEVIATION	DIFFERENCE IN PERCENTAGE DEVIATION
	Base of prediction	Actual deviation	Percentage deviation	Base of prediction	Actual deviation	Percentage deviation		
Feb. to Oct.	November	35.01	25.20	Nov.–Jan.	33.84	24.36	+1.17	+0.84
Feb. to Oct.	December	34.94	25.15	Nov.–Jan.	33.84	24.36	+1.10	+0.79
Feb. to Oct.	January	34.61	24.91	Nov.–Jan.	33.84	24.36	+0.77	+0.55
Mar. to Oct.	December	32.74	25.38	Dec.–Feb.	31.35	24.30	+1.39	+1.08
Mar. to Oct.	January	32.52	25.22	Dec.–Feb.	31.35	24.30	+1.17	+0.92
Mar. to Oct.	February	30.86	23.92	Dec.–Feb.	31.35	24.30	−0.49	−0.38
Apr. to Oct.	January	30.30	27.04	Jan.–Mar.	28.62	25.54	+1.68	+1.50
Apr. to Oct.	February	29.17	26.03	Jan.–Mar.	28.62	25.54	+0.55	+0.49
Apr. to Oct.	March	28.40	25.34	Jan.–Mar.	28.62	25.54	−0.22	−0.20
May to Oct.	February	27.10	28.66	Feb.–Apr.	26.14	27.64	+0.96	+1.02
May to Oct.	March	27.09	28.65	Feb.–Apr.	26.14	27.64	+0.95	+1.01
May to Oct.	April	26.50	28.02	Feb.–Apr.	26.14	27.64	+0.36	+0.38
June to Oct.	March	25.49	35.05	Mar.–May	24.70	33.96	+0.79	+1.09
June to Oct.	April	25.02	34.40	Mar.–May	24.70	33.96	+0.32	+0.44
June to Oct.	May	24.81	34.11	Mar.–May	24.70	33.96	+0.11	+0.15
July to Oct.	April	22.51	42.05	Apr.–June	21.42	40.01	+1.09	+2.04
July to Oct.	May	22.34	41.72	Apr.–June	21.42	40.01	+0.92	+1.71
July to Oct.	June	21.44	40.04	Apr.–June	21.42	40.01	+0.02	+0.03
Aug. to Oct.	May	18.28	51.38	May–July	17.05	47.93	+1.23	+3.45
Aug. to Oct.	June	18.18	51.09	May–July	17.05	47.93	+1.13	+3.16
Aug. to Oct.	July	17.14	48.17	May–July	17.05	47.93	+0.09	+0.24
Sept. to Oct.	June	13.48	71.47	June–Aug.	12.52	66.38	+0.96	+5.09
Sept. to Oct.	July	13.20	70.00	June–Aug.	12.52	66.38	+0.68	+3.62
Sept. to Oct.	August	12.12	64.26	June–Aug.	12.52	66.38	−0.40	−2.12
October	July	7.06	120.23	July–Sept.	6.44	109.67	+0.62	+10.56
October	August	6.92	117.84	July–Sept.	6.44	109.67	+0.48	+8.17
October	September	5.71	97.27	July–Sept.	6.44	109.67	−0.73	−12.40

is not so satisfactory as might be desired. Great irregularities are to be expected when a flock of only 415 birds is divided up into such a large number of classes. The results indicate that if applied to larger flocks the prediction equations for the production of a group of remaining months from three-months recorded production might be made with greater precision.

In comparing the error of prediction for a group of subsequent months from three-month periods with the errors of prediction for the same period when prediction is made from single months we require three sets of equations for the prediction of the yields of a group of months from a monthly record. Two sets of these have been given on page 282 and used in comparison with the results of prediction from bimonthly periods.

The additional equations required are:

Month from which prediction is made	Period for which prediction is made	Prediction equation
November	February to October	$E_9 = +134.546 + 1.143\,e_1$
December	March to October	$E_8 = +122.721 + 1.165\,e_2$
January	April to October	$E_7 = +106.659 + 1.033\,e_3$
February	May to October	$E_6 = +80.512 + 1.342\,e_4$
March	June to October	$E_5 = +47.994 + 1.462\,e_5$
April	July to October	$E_4 = +30.111 + 1.219\,e_6$
May	August to October	$E_3 = +13.094 + 1.008\,e_7$
June	September to October	$E_2 = +3.462 + 0.649\,e_8$
July	October	$E_1 = +0.297 + 0.240\,e_9$

When we compare the results for the prediction of the yield of a group of subsequent months from single monthly records and from trimonthly records of production we find that the differences in errors of prediction are surprisingly small. Specifically we note that in the case of the average deviation with regard to sign, shown in the two last columns of table 11, the differences in actual errors range from 0.03 to 3.42 eggs while the differences in percentage values range from 0.05 to 10.55. In some cases the three-month period gives a numerically larger error of prediction while in other cases the one-month period gives the larger error.

When the comparison is made on the basis of average deviation without regard to sign (table 12) the single-month period gives a slightly larger average deviation in most cases, 23 out of 27 cases. The differences are, however, very small, varying from 0.12 to 1.14 eggs.

Similar results are obtained when the comparisons (between the single component months and the three-months record as bases of prediction) are based upon square root of mean square deviation (table 13). In 24 of the 27 cases prediction from a monthly record gives slightly more variable

rrors than prediction from the combined record of three months. The
differences are, however, insignificant, varying from 0.02 to 1.68 eggs. It
s clear, therefore, that if the linear equation be used for the purpose of
predicting the yield of a group of remaining months, about as good results
or practical purposes may be obtained from single month records as from
he sum of three months records.

It is quite possible that with equations other than the linear this will
ot be the case. Such equations will be investigated in future work.

Comparison of the two- and three-month periods as bases for the prediction of the egg record of the subsequent months

In the foregoing discussion comparisons between the value of single-
month periods and two-month periods and between single-month and three-
month periods as bases for prediction have been made. It will be of some
interest to compare two- and three-month periods in the same way. Cer-
ain of the data may be rearranged from preceding tables. Special calcu-
ations would, however, be necessary to complete all of the possible com-
parisons. It is evident that for a critical comparison between the two
roups it is necessary to deal with the egg record of a group of remaining
months. Thus in comparing November-to-January production with
November-and-December or December-and-January production as bases
f prediction it is necessary to determine the accuracy with which the egg
production of February to October may be predicted since none of the
months included in the base of prediction should also occur in the period
or which prediction is made.

Limiting our attention to the comparisons which can be made from the
ata in the preceding tables[7] we note that in some cases there is a larger
verage deviation with regard to sign in predicting from two-months and
n some cases a larger error in predicting from three-months production.

The same may be shown to be true for the average deviation without
egard to sign and for the square root of mean square deviation of the
predicted from the actual values. Thus there is little practical advantage
n dealing with three-months production as compared with two-months
production as bases for the prediction of the record of a group of subsequent
months.

[7] The subsidiary tables upon which the following conclusions were based may be formed from
ables 5 to 7 and 11 to 13. It seems unnecessary to publish these tables here.

Comparison of the four periods as bases for the prediction of the egg record of the year

In the introductory sections of this paper we called attention to the so-called periods or cycles of egg production which have been recognized by a number of students of fecundity in the domestic fowl. It might at first seem desirable to compare the results of predicting from these periods.

Since these periods are consecutive and together make up the entire laying year it is impossible to obtain any common basis for testing their efficiency such as has been found in periods of subsequent months in preceding tests.

In view of this fact it does not seem desirable in this place to go into the question of the comparison of these conventional periods as bases of prediction. Practically all of the data required for such comparison as can be made appear in the foregoing tables 2 to 13. The reader who desires to do so may abstract the constants from these tables.

SUMMARY AND CONCLUSIONS

The specific purpose of the present paper, which is one of a series dealing with the general problem of variation and correlation of egg production in the domestic fowl, is to consider the possibility of predicting the future egg production or the total annual egg production of White Leghorn birds from the record of an individual month or a group of consecutive months.

The investigation has been carried out because of two convictions: First, factors underlying the distribution, inheritance and interrelationships of fecundity in birds present a problem of first-rate biological importance. Second, that it is one of the functions of the biologist to provide the agricultural economist with the quantitative constants and formulae upon which the scientific agriculture of the future must largely rest.

The method followed has been to determine a series of prediction equations based on the experience of six years (1911 to 1917) of the INTERNATIONAL EGG-LAYING CONTEST at Storrs and to test these equations upon an additional series of 415 birds studied at Storrs in 1917–1918. Thus the equations have been tested upon a different series of birds from that upon which they were based, but upon birds maintained under conditions comparable with those upon whose record the fundamental equations were based.

The results show that the annual egg record of a series of birds may be predicted with a reasonably high degree of accuracy when their performance for a single month is known. Somewhat higher accuracy may be obtained

hen the record of two or more months is taken into consideration, but the nprovement due to an increase in the number of months upon which rediction is based is not great.

Prediction of the egg record which will be made by groups of birds sub-:quently to the month or group of months chosen as a basis of prediction in also be made, but the accuiacy of prediction decreases rapidly as the eriod for which prediction is made becomes shorter.

The results show that in the case of a flock of White Leghorn fowl, which essentially identical in genetic composition and maintained under essen-ally uniform conditions from year to year, it is quite possible to estimate nnual egg production from the record of either a single month or of two r three consecutive months with a high degree of accuracy. The same presumably true of other breeds as well. This point is now under ivestigation.

It is probably not feasible to use the equations given in this paper for ocks differing greatly in genetic composition or in conditions of mainte-ance from that upon which these equations were based. The problem the determination of corrective terms by which the equations may be pplied to flocks other than that upon which they are based is now under ivestigation.

LITERATURE CITED

_DER, B., and EGBERT, A. D., 1918 A quick method of obtaining individual egg records without the trap nest. Bull. Utah Agric. Exp. Sta. 162.

_AKESLEE, A. F., HARRIS, J. ARTHUR, WARNER, D. E., and KIRKPATRICK, W. F., 1917 Pigmentation and other criteria for the selection of laying hens. Bull. Storrs Agric. Exp. Sta. 92, pp. 95–194.

\RD, L. E., 1917 A study of egg production in the White Leghorns. Bull. Storrs Agric. Exp. Sta. 91, pp. 41–90.

\RRIS, J. ARTHUR, BLAKESLEE, A. F., and WARNER, D. E., 1917 Body pigmentation and egg production in the fowl. Proc. Nation. Acad. Sci. 3: 237–241. _Fig. 1–3._

\RRIS, J. ARTHUR, BLAKESLEE, A. F., and KIRKPATRICK, W. F., 1917 Interperiodic correlation in the egg production of the domestic fowl. Proc. Nation. Acad. Sci. 3: 565–569. _Fig. 1–2._

\RRIS, J. ARTHUR. BLAKESLEE, A. F., and KIRKPATRICK, W. F., 1918 The correlation between egg production during various periods of the year in the domestic fowl. Genetics 3: 27–72. _Fig. 1–11._

\RRIS, J. ARTHUR, BLAKESLEE, A. F., WARNER, D. E., and KIRKPATRICK, W. F., 1917 The correlation between body pigmentation and egg production in the domestic fowl. Genetics 2: 36–77. _Fig. 1–16._

THE PREDICTION OF ANNUAL EGG PRODUCTION
FROM THE RECORDS OF LIMITED PERIODS

BY

J. ARTHUR HARRIS, W. F. KIRKPATRICK AND A. F. BLAKESLEE

Station for Experimental Evolution, Carnegie Institution, and the Storrs Agricultural Experimental Station

THE PREDICTION OF ANNUAL EGG PRODUCTION FROM THE RECORDS OF LIMITED PERIODS

By J. ARTHUR HARRIS, W. F. KIRKPATRICK AND A. F. BLAKESLEE

STATION FOR EXPERIMENTAL EVOLUTION, CARNEGIE INSTITUTION, AND THE STORRS
AGRICULTURAL EXPERIMENTAL STATION

For the past several years the writers have been considering the possibility of predicting the annual egg production of the domestic fowl from the records of short periods of time. Such records may be determined by trap-nesting, or by the use of other criteria when the maternity of the eggs is not required for breeding purposes.[1]

The first definite step in the direction of the use of the egg record of a short period for the prediction of the production during a subsequent or a longer period was, as far as we are aware, taken in 1917 when it was shown[2] that in a heterogeneous series of birds such as are submitted by practical breeders in egg laying contests, the October egg production is correlated with that of every other month of the year. The investigation was carried much further in a second memoir[3] in which the correlations between the records of the individual months and the production of the whole year, between the records of the individual months and those of the remaining 11 months of the year, and between the production of 5 of the individual months and the production of all the other individual months, were published for two series of birds. In this paper the equations for the prediction of total annual production from the record of the individual months were given.

Our purpose here is to state briefly the results of a first test of the possibility of utilizing the linear regression equation (which is strictly valid only for the population from which it is deduced) for the prediction of the records of the birds of a flock the performance of which is unknown as far as the determination of the constants of the equations is concerned.

In a population the straight line relating the egg production of a period E_p with that of a period used as a basis of prediction, e, is

$$E_p = \left(\bar{E}_p - {}^r e E_p \frac{\sigma E_p}{\sigma_e} \; \bar{e} \right) + {}^r e E_p \frac{\sigma E_p}{\sigma_e} \; e$$

where the bars denote means, the sigmas represent standard deviations, and r indicates the correlation of the two variables in the standard population.

The value of E_p given by the equation is the theoretical mean production for the array of individuals of any class with respect to e. The assumption to be tested is that we may write $E_p{}'$ and e' instead of E_p and e, where $E_p{}'$ is the theoretical mean production for a period p of the array of birds of any grade of production e in the period used as a basis of prediction in a series of birds which are not involved in the data upon which the equations were based, but to which the equations are to be applied for practical purposes.

The essential practical requisites for such prediction equations are: (1) That the errors of prediction shall be distributed about the true numbers in such a manner that estimation will not in the long run be either too high or too low. (2) That the magnitude of the deviation of the predicted from the observed egg production shall be as small as possible.

Let $E_p{}''$ be the actual and $E_p{}'$ the predicted egg production of an individual bird for any period, p, in a flock to which the equation is being applied. The error of prediction is then $E_p{}' - E_p{}''$. The average of these errors, with regard to sign

$$\frac{1}{N} \Sigma \left(E_p{}' - E_p{}'' \right)$$

furnishes a measure of the success with which the first requirement, (1) above, is met. The average of these errors without regard to sign furnishes a measure of the average error above or below the true production of the individual birds of a flock. The square root of mean square deviation

$$\left[\frac{1}{N} \Sigma \left(E_p{}' - E_p{}'' \right)^2 \right]^{1/2}$$

furnishes a measure of this error which weights larger errors.

The errors may be expressed in actual numbers of eggs, or, in relative terms, as percentages of the mean production of the period and flock for which prediction is made. Both methods have been used in testing the equations.

In testing the efficiency of such equations for purposes of prediction we have proceeded in a purely objective manner. Working on the assumption that the crucial test of any theory is its capacity for predicting

the unknown, we have calculated equations based upon the data of the International Egg Laying Contest at Storrs, Conn., during the six contest years, 1911–1917, inclusive. We have then used these equations to predict the annual production (and the production of groups of months) for the birds of the 1917–'18 contest, using as a basis of prediction the individual months of the laying year separately, pairs of successive months and groups of three months. Our conclusions concerning the value of the equations depend, therefore, not upon *a priori* considerations but upon the results of actual tests of accuracy of prediction for series which were unknown as far as the determination of the constants of the equations is concerned.

Consider first of all the results of the attempts to predict the annual egg production of 415 White Leghorn birds observed at Storrs from Nov. 1, 1917 to Oct. 31, 1918 from the records of a single month's production.

The results of the three criteria of accuracy of prediction are summarized in table 1.

TABLE 1

ERRORS OF PREDICTION OF ANNUAL EGG PRODUCTION FROM THE RECORDS
OF INDIVIDUAL MONTHS

MONTH USED AS BASE OF PREDICTION	AVERAGE DEVIATION WITH REGARD TO SIGN		AVERAGE DEVIATION WITHOUT REGARD TO SIGN		SQUARE ROOT OF MEAN SQUARE DEVIATION	
	Actual deviation	Percentage deviation	Actual deviation	Percentage deviation	Actual deviation	Percentage deviation
November	+ 2.39	1.52	29.59	18.78	38.65	24.52
December	− 0.49	0.31	29.26	18.57	37.61	23.86
January	+ 2.58	1.64	30.09	19.09	38.77	24.60
February	+ 0.06	0.04	27.28	17.31	34.70	22.02
March	− 1.63	1.03	27.95	17.73	34.28	21.75
April	− 6.23	3.95	28.72	18.22	35.31	22.40
May	+ 7.02	4.45	28.62	18.16	35.89	22.77
June	− 5.21	3.31	29.03	18.42	36.53	23.18
July	− 5.27	3.34	28.35	17.99	35.89	22.77
August	− 0.82	0.52	26.87	17.05	34.34	21.79
September	+ 4.78	3.03	24.78	15.72	32.94	20.90
October	+ 3.95	2.51	27.37	17.37	36.47	23.14

Considering first of all the absolute values we note that the average errors with regard to sign are generally low. Thus the prediction from November and from January production gives on the average 3 eggs too many for the year. For December, February, March and August the prediction is in error by less than 2 eggs. The values predicted from April, May, June, July, September and October records are from 4 to 7 eggs in error.

The average deviations without regard to sign are of course much larger since they constitute a measure of the error of prediction of the records of individual birds. They range from 24.8 to 30.1 eggs. The significance of errors of this magnitude will be more clearly brought out later.

The square root of mean square deviation also shows considerable regu-

larity from month to month.   These measures are naturally considerably larger than the average deviation without regard to sign.   They range from 32.9 to 38.8 eggs.

It is clear that the annual egg production of birds similar in origin to the series upon which the prediction equations were based and maintained under similar conditions may be predicted with a relatively high degree of accuracy providing their record for any month is definitely known.

The order of the errors will be more readily understood by expressing them in relation to the average production of the flock, as shown by the percentage deviations.

We note that in predicting from December, February and August records the average error with regard to sign *is less than one per cent of the average annual yield of the flock.*   In predicting from November, January and March the error lies between one and two per cent.   When April, May, June, July, September and October records are used as a basis of prediction the average errors of prediction are from 2.50 to 4.50 per cent of the average annual yield.

The average deviations without regard to sign are less than 20 per cent of the annual production.   The values for the individual months range from 15.7 for September to 19.1 for January.

The square root of mean square deviation is less than 25 per cent of the average annual production.   The individual values range from 20.9 for September to 24.6 for January.

These two latter tests may at first seem to indicate very unsatisfactory prediction.   Such is not, however, the case.   These give the average errors *either above or below the true record made in the prediction of the results for an individual bird.*   The thing which is required in practise is generally *the prediction for a group of birds of a particular grade* of egg record for the month used as a base of prediction.   In a flock of 415 birds this has been shown to be possibe *with an error of less than 5 per cent of the annual production* when prediction is made from the record of *any* month of the year; and *with an error of less than 1 per cent when prediction is based upon the records of a number of the individual months.*

Lack of space precludes a discussion of the results of the prediction of the annual record of the bird from the combined record of two consecutive months.   We may, however, illustrate the accuracy of prediction from the combined record of two consecutive months by means of the figures in diagram 1 which shows the accuracy of prediction from November plus December and from April plus May in comparison with the results of prediction from November and April.   In these the estimated production is shown by a straight line.

The actual production for the year for which prediction is made is shown by solid dots for each group of birds as classified by monthly or bimonthly

DIAGRAM 1

ɛcord. The shaded areas are determined as follows. The birds were
rst grouped into classes of five eggs range with respect to number of eggs
ιid during the period of time used as a basis of prediction. The birds of
ιese classes of five eggs range were further subdivided into those in which
ɛtual egg production was greater than the predicted and those in which
ιe actual number was less than the predicted number.[4] The average

error of prediction was determined for each of these groups, and these averages represent the upper and lower limits of the shaded areas. The upper limit represents, therefore, the average deviation (for the period for which prediction is made) of all birds which make a higher record than that predicted for their class. The lower limit of the shaded area marks the average deviation for all birds which show an egg record lower than that predicted. These diagrams, which are quite typical of the whole series, certainly indicate excellent prediction.

The results for the combined records of three consecutive months are shown in table 2. These show that greater accuracy of prediction may be obtained when the records of three months are used as a basis of prediction. Such a result is to be expected on *a priori* grounds. A careful comparison of the constants in tables 1 and 2 will show, however, that the improvement resulting from the trebling of the number of months used as a basis of prediction is not great.

TABLE 2

ERRORS OF PREDICTION OF ANNUAL EGG PRODUCTION FROM THE RECORD
OF THREE CONSECUTIVE MONTHS

THREE MONTHS USED AS BASE OF PREDICTION	AVERAGE DEVIATION WITH REGARD TO SIGN		AVERAGE DEVIATION WITHOUT REGARD TO SIGN		SQUARE ROOT OF MEAN SQUARE DEVIATION	
	Actual deviation	Percentage deviation	Actual deviation	Percentage deviation	Actual deviation	Percentage deviation
Nov.–Jan.	+ 2.09	1.33	25.93	16.45	33.84	21.47
Dec.–Feb.	+ 0.78	0.49	25.31	16.06	32.65	20.72
Jan.–Mar.	+ 0.49	0.31	25.29	16.05	31.58	20.04
Feb.–Apr.	— 4.07	2.58	24.16	15.33	29.77	18.89
Mar.–May	— 0.73	0.46	25.42	16.13	31.14	19.76
Apr.–June	— 2.31	1.47	24.33	15.44	30.59	19.41
May–July	— 2.12	1.35	24.20	15.36	29.40	18.65
June–Aug.	— 5.35	3.39	23.49	14.90	29.80	18.91
July–Sept.	— 0.20	0.13	21.36	13.55	28.10	17.83
Aug.–Oct.	+ 3.91	2.48	21.59	13.70	29.23	18.55

Prediction of the number of eggs which will be laid in the period subsequent to the month or group of months used as a basis of prediction may also be made. The errors for such a series of predictions, in which each individual month of the year (with the exception of the final month) has served as a basis for the prediction of the egg production of the remaining months of the year, are shown in table 3. The constants in this table show that when the period for which prediction is made is a long one a degree of accuracy fairly comparable with that for the whole year is attainable. The absolute values of the average deviation without regard to sign and of the square root of mean square deviation necessarily become smaller as the period for which prediction is made becomes shorter. The relative (percentage) error, however, increases. Thus the accuracy of prediction decreases rapidly as the period for which prediction is made becomes shorter.

TABLE 3

Errors of Prediction of the Record of a Period of Months from the Record of Individual Preceding Months

PERIOD FOR WHICH PREDICTION IS MADE	MONTH USED AS BASE OF PREDICTION	AVERAGE DEVIATION WITH REGARD TO SIGN		AVERAGE DEVIATION WITHOUT REGARD TO SIGN		SQ. ROOT OF MEAN SQ. DEVIATION	
		Actual deviation	Percentage deviation	Actual deviation	Percentage deviation	Actual deviation	Percentage deviation
Dec.–Oct.	November	+ 2.39	1.57	29.59	19.49	38.65	25.46
Jan.–Oct.	December	+ 0.24	0.16	28.43	19.53	36.63	25.16
Feb.–Oct.	January	+ 2.37	1.71	26.48	19.06	34.61	24.91
Mar.–Oct.	February	+ 0.09	0.77	24.07	18.66	30.86	23.92
Apr.–Oct.	March	− 0.24	0.21	22.81	20.36	28.40	25.34
May–Oct.	April	− 4.00	4.23	21.36	22.59	26.50	28.02
June–Oct.	May	+ 3.62	4.98	19.89	27.35	24.81	34.11
July–Oct.	June	− 2.86	5.34	17.62	32.91	21.44	40.04
Aug.–Oct.	July	− 3.71	10.43	13.91	39.09	17.14	48.17
Sept.–Oct.	August	− 2.56	13.57	9.76	51.75	12.12	64.26
October	September	− 0.45	7.67	4.57	77.85	5.71	97.27

The results of this investigation, taken as a whole, show that in the case f a flock of White Leghorn fowl which is essentially identical in genetic omposition and maintained under essentially uniform conditions from ear to year it is quite possible to estimate annual egg production from the ecord of either a single month or of two or three consecutive months with high degree of accuracy. The same is presumably true of other breeds s well. This point is now under investigation.

It is not possible to use the equations given in this paper for flocks differ-1g greatly in genetic composition or in conditions of maintenance from 1at upon which these equations were based. The problem of the deter-1ination of corrective terms to be used when the equations are applied to ocks other than that upon which they are based is now under investiga-.on.

A detailed account of these investigations is now in press in *Genetics*.

[1] Alder and Egbert, *Bull. Utah Agr. Exp. Sta.*, No. 162, 1918.

[2] Harris, Blakeslee and Warner, These Proceedings 3, 1917 (337–341); Harris, lakeslee, Warner and Kirkpatrick, *Genetics*, 2, 1917 (36–77).

[3] Harris, Blakeslee and Kirkpatrick, These Proceedings 3, 1917 (565–569); *enetics*, 3, 1918 (27–72).

[4] A range of five eggs was used in order to obtain a number of birds sufficiently large ₃ reduce somewhat the irregularities due to the errors of random sampling. The ⁻rors of prediction were in each case determined for classes of unit range. Grouping used for graphic representation merely. The average deviations represented by the ₁it of the shaded zone are to be thought of as measured from a line perpendicular ₃ the ordinates and intersecting the prediction line on the mid-ordinate of the 5-egg ass.

A BIOMETRIC STUDY OF HUMAN BASAL METABOLISM

BY

J. ARTHUR HARRIS AND FRANCIS G. BENEDICT

Nutrition Laboratory and Station for Experimental Evolution, Carnegie Institute of Washington

A BIOMETRIC STUDY OF HUMAN BASAL METABOLISM

By J. Arthur Harris and Francis G. Benedict

Nutrition Laboratory and Station for Experimental Evolution,
Carnegie Institution of Washington

Communicated October 8, 1918

Investigators are now generally agreed that the metabolism, expressed in terms of calories per unit of time, of the normal subject shall be taken as a basis of comparison in the investigation of all the special problems of human nutrition, for example, that of the requirements for muscular activity, that of the influence of specific diseases or of the level of nutrition upon metabolism, that of the change of metabolic activity with age, and so forth. Critical investigations in both European and American laboratories have shown that the gaseous metabolism is so affected by various factors that determinations which are to serve as a standard must be made under very exactly controlled conditions. It is not merely necessary to devise apparatus in which the physical difficulties of direct calorimetry (or of the exact measurement of gaseous exchange from which heat production may be computed) are overcome. Certain biological factors must be ruled out. Those of greatest importance as sources of experimental error are muscular activity and the stimulatory action of recently ingested food. The heat production of the individual in a state of complete muscular repose 12–14 hours after the last meal, i.e., in the postabsorptive condition, has been called the basal metabolism.

For a decade the Nutrition Laboratory has been engaged in carrying out a series of determinations of basal metabolism in normal human individuals of both sexes and of widely different ages. These have been made with all the modern refinements of method and manipulation. The subjects were in presumably good health. All those with febrile temperature were discarded. All were in the postabsorptive condition. Perfect muscular repose during the short periods required for indirect calorimetry was assured by an automatic record of all movements, even those imperceptible to a trained observer.

Measurements on 136 men, 103 women and 94 new-born infants have been analyzed biometrically with the purpose of determining the statistical constants (means, standard deviations, coefficients of variation, coefficients of correlation, and regression equations) which may serve as standard constants in work on human metabolism until those based on more extensive series of data are available. In carrying out this analysis we have proceeded on the conviction that the widest possible usefulness of laboratory investigations of human metabolism will result from basing measurements upon individuals who are in presumably good health, but who are otherwise typical of the population at large. It is only when the subjects used for experimentation are representative of the general population in type, variability and correlation that results of laboratory research upon limited series of individuals may be safely generalized for rationing or for other practical social applications. Statistical tests of the suitability of the series used in the present investigation have been applied.

The relationship between certain of the physical and physiological measurements of the individual and between the various physiological measurements has been determined. Our series of data show practically no relationship between basal or minimum pulse rate and stature or body weight in adults. There is a low but significant positive correlation between minimum or basal pulse rate and gaseous exchange and heat production. The Nutrition Laboratory has long emphasized the correlation between pulse rate and metabolism in the same individual, that is, the intra-individual correlation between the rate of the heart beat and the amount of the katabolism. Here, however, we are dealing with the problem of the relationship between the minimum or basal pulse rate of a series of individuals and their basal metabolism—that is, with inter-individual correlation.

There is a substantial correlation between stature and heat production. The correlation between body weight and heat production is higher being of the order $r = 0.75$ to $r = 0.80$ in the new-born infants, of the order $r = 0.80$ in men and $r = 0.60$ in women. Analysis by means of partial correlation coefficients indicates that both stature and body weight have independent significance as bases for the prediction of the basal metabolism.

The change in basal metabolism with age during the period of adult life has been shown to be well represented by the linear equations,

For men ($N = 136$)

$$h = 1823.80 - 7.15, a, \quad h_k = 28.703 - 0.112\, a, \quad h_d = 1022.17 - 3.60\, a.$$

For women ($N = 103$)

$$h = 1420.47 - 2.29\, a, \quad h_k = 28.308 - 0.124\, a, \quad h_d = 942.25 - 2.96\, a.$$

where h = total heat production in calories per 24 hours, h_k = calories per kilogram of body weight, h_d = calories per square meter of body surface as estimated by the Du Bois height-weight chart. Thus in men the daily heat

production decreases about 7.15 while in women it decreases about 2.29 calories per year. Women are smaller than men and have a lower heat production. When the decrease in metabolism with age is expressed in calories per kilogram of body weight or in calories per square meter of body surface, the results for the two sexes are much more nearly identical.

The problem of the difference in the metabolism of men and women, dealt with in the past by a number of writers, has been reconsidered on the basis of the larger series of data now available. The average daily (24 hours) basal heat production of men is 1632 calories whereas that of women is 1349 calories. Thus women have an average daily heat production about 300 calories less than that of men. But women are smaller than men. If correction for body size be made by expressing heat production in calories per kilogram of body weight, it is 25.7 calories in the 136 men as compared with 24.5 calories, or 1.2 calories per kilogram less, in the 103 women. On the basis of heat production per square meter of body surface as estimated by the Du Bois height-weight chart the men show an average daily heat production of 925 calories as compared with 850 calories, or 75 calories less, in the 103 women. The most critical test of the difference of men and women in the level of metabolism is that furnished by a modification and extension of the selected group method of Benedict and Emmes. In the new method the control values for the several groups of women are not the empirical constants for men of as nearly as possible like stature and body weight but are determined from equations taking into account stature, weight and age in all the available data for men. Analysis shows that, however expressed, the metabolism of American women is lower than that of the men. Our results show that the differentiation of the sexes is not evident in infancy. They do not confirm the conclusion of Sondén and Tigerstedt that the difference between men and women tends to disappear with age. Instead we find the difference in the metabolism of men and women well-marked throughout the period of adult life.

The validity of the so-called body surface law, according to which metabolism is proportional to the superficial area of the individual, i.e.,

$$h = a\bar{h}_a$$

where a = superficial area and $\bar{h}_a$ = mean heat production per unit of time per square meter of body surface in a standard series, has been critically tested. It has been shown that the supposed proofs of its validity hitherto adduced are erroneous. Heat production is not 'proportional to body surface but not to body weight' as has been asserted to be the case, but is highly and about equally correlated with both body weight and body surface. It has been shown that as a basis for predicting the heat production of a subject the above relationship is less satisfactory than multiple regression equations involving stature, weight and age. Thus the 'body surface law' is deprived of its unique significance as a basis for the prediction of the me-

abolism of an unknown subject. An analysis of the data of actual experimentation on subjects at changing levels of nutrition shows that the changes in metabolism are not proportional to those in body surface. Surface area may not be looked upon as a determining factor in basal metabolism.

The closest prediction of the daily heat production of a subject can be made by the use of the multiple regression equations,

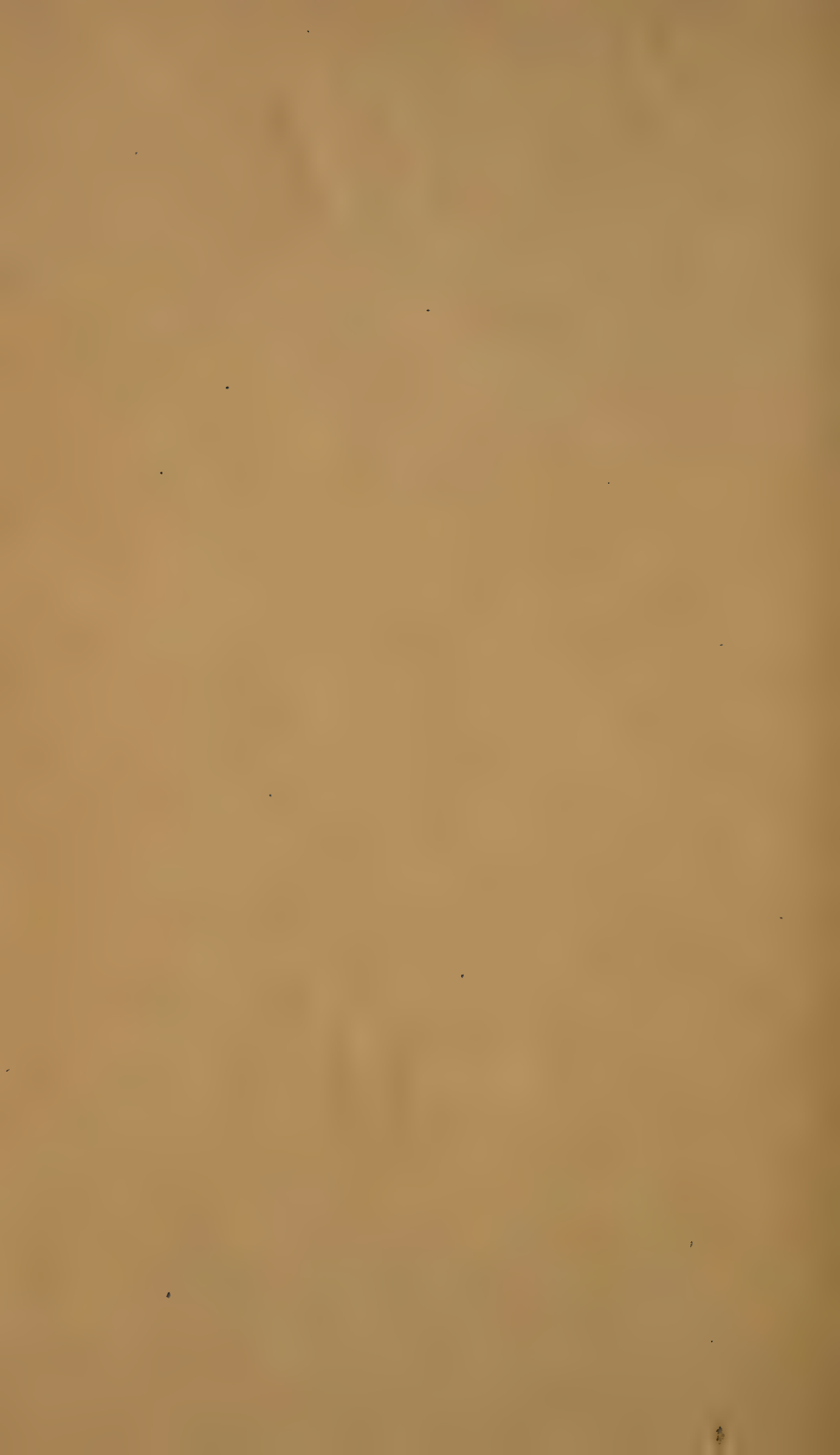For men, $h = 66.4730 + 13.7516\,w + 5.0033\,s - 6.7550\,a$

For women, $h = 655.0955 + 9.5634\,w + 1.8496\,s - 4.6756\,a$

where h = total heat production per 24 hours, w = weight in kilograms, s = stature in centimeters, and a = age in years. These equations have been tabulated for values of weight from 25.0 to 124.9 kgm., for stature from 151 to 200 cm., and for age from 21 to 70 years, so that the most probable basal metabolism of an unknown subject may be easily determined.

Such tables should render service in clinical and other fields of applied calorimetry. Their usefulness has been demonstrated in testing the typical or atypical nature of series of metabolism measurements, in investigating the differentiation of the sexes with respect to metabolic activity, of the metabolism of athletes as compared with non-athletic individuals, and of individuals suffering from disease.

The detailed measurements and statistical constants, with full discussions of pertinent literature, are about to appear in Publication No. 279 of the Carnegie Institution of Washington.

BIOMETRIC STANDARDS FOR ENERGY REQUIREMENTS IN HUMAN NUTRITION

By Dr. J. ARTHUR HARRIS and Dr. FRANCIS G. BENEDICT

BIOMETRIC STANDARDS FOR ENERGY REQUIREMENTS IN HUMAN NUTRITION

By Dr. J. ARTHUR HARRIS and Dr. FRANCIS G. BENEDICT

CARNEGIE INSTITUTION OF WASHINGTON

ONE of the primary requisites in all of the exact sciences is the establishment of standard bases of comparison. For a decade the Nutrition Laboratory of the Carnegie Institution of Washington has been engaged in the precise investigations which must underlie the establishment of such standards in human nutrition.

This is an undertaking of the greatest practical importance. In times of peace, industrial efficiency and the physical well-being of the population demand exact knowledge of the amount and proportion of the different kinds of food which should be taken by the individual. If communities or nations are to be stringently rationed during periods of emergency, it is also necessary to know the minimum amounts of food required to maintain health and efficiency.

The problem is also one of great complexity. Aside from all questions concerning the chemical composition, digestibility and other physiological properties of the various foods, there are a large number of problems concerning the characteristics of human individuals which must be taken into account.

For example, it is obvious that those who are engaged in severe muscular work must consume larger quantities of food supplying energy than those who are less active. It might seem reasonable to suppose that larger individuals would re-quire more food to carry on their normal activities than those who are physically smaller. It is a matter of common observation that older men and women demand smaller rations than those in the earlier stages of life.

All these questions require precise investigation before one is justified in drawing conclusions concerning them. If such investigations are to be used as a basis of recommendation concerning diet in peace or of regulation of diet in war, it is essential that the laws of energy transformation be expressed in a quantitative form.

Nutritional physiologists agree that, as far as energy is concerned, the food requirements of the living organism shall be expressed in calories per unit of time. Thus a physical standard is taken over from the quantitative sciences of physics and chemistry. Theoretically, then, the metabolism must be determined by placing the subject in a calorimeter and directly measuring the number of calories produced. This has been done in a large number of cases.

Since, however, the setting free of energy in the human body is merely a process of combustion, the measurement of the amount of oxygen consumed and the quantity of carbon dioxide excreted from the lungs should furnish a good index of heat production. Thus the nutritional physiologist may avail himself of the method of *indirect calorimetry* as well as of *direct calorimetry*. Heat production, in short, may be determined in a calorimeter or it may be computed from the gaseous exchange as measured in a respiration chamber.

The development of apparatus by which the heat production of the living organism may be directly measured in the calorimeter or by which the gaseous exchange may be precisely determined in the respiration chamber has occupied the attention of a large number of ingenious experimenters, among whom may be mentioned Lavoisier, Rubner, Zuntz, Atwater, Rosa, Lusk and Du Bois. The labors of these and others have brought the apparatus for the measurement of both heat production and gaseous exchange to such a high degree of refinement that the manipulative phases of nutritional physiology may be regarded as among the most exact techniques of biological experimentation. Extensive comparative studies have shown that, in the case of human subjects, it is much simpler and essentially as accurate to calculate the heat production indirectly from the gaseous exchange than to measure it in the calorimeter.

The problem is not, however, solely one of physical and chemical measurement. A number of biological factors must be taken into account. Muscular activity and the stimulatory action of recently ingested food are of chief importance. The apparatus with which students of human nutrition now work

has been brought to such a stage of perfection as to measure the energy transformation accompanying such slight muscular activity as that required in the raising of the hand from the side to the mouth. The cost in calories of masticating food may be directly measured. For example, recent studies at the Nutrition Laboratory by Carpenter and Benedict have shown that the muscular work in chewing gum may increase heat production approximately 17 per cent. The difference between the heat production of a new-born infant asleep in its basket and crying can be precisely measured. Thus Talbot and Benedict found that the metabolism of the new-born infant is increased on the average by 65 per cent. in crying with its attendant muscular activity. Students of animal nutrition have long realized that the demands for energy of an animal standing are far higher than that of the same beast lying down. This fact must be taken into account in computing the maintenance requirements of cattle and other domestic animals.

Heat production is greatly increased after eating, and the amount of the increase is closely dependent on the nature of the food consumed. For example, the metabolism of a subject may be increased by 25 per cent. after a meal consisting chiefly of carbohydrates, but by as much as 45 per cent. after a heavy protein meal.

It is necessary, therefore, to eliminate all such factors in determining the standards which shall serve as bases of comparison in applied nutritional physiology.

Since the outbreak of the war, and particularly since our own participation in the conflict, the Nutrition Laboratory has, in addition to extensive investigations on the influence of subnormal rationing upon health and efficiency, pushed forward as rapidly as possible its work on the establishment of nutritional standards. One phase of this program has been the statistical investigation of the so-called basal metabolism of the human individual.[1]

Physiologists have gradually come to a general agreement that the heat production at complete muscular repose and in the post-absorptive state—*i. e.*, about twelve hours after the last meal—shall be called the basal metabolism and shall be used as a standard of comparison in the investigation of all the special problems of human nutrition.

[1] The detailed measurements and the statistical constants, with full discussions of pertinent literature, are about to appear in Publication 279 of the Carnegie Institution of Washington. We shall not, therefore, burden this brief outline with references to literature or statistical detail. Two of the diagrams used here are redrawn from this publication.

For several years the Nutrition Laboratory has been engaged in the measurement of basal metabolism in normal human individuals of both sexes and of widely different ages. These have been made with all the modern refinements of method and manipulation. The subjects were in presumably good health. All those with febrile temperature were rejected. All were in the post-absorptive condition. Perfect muscular repose during the short periods required for indirect calorimetry was assured by instruments providing an automatic record of all movements, even those imperceptible to trained observers.

FIG. 1. FREQUENCIES OF MEN AND WOMEN PRODUCING VARIOUS NUMBERS OF CALORIES PER TWENTY-FOUR HOURS.

In carrying out a biometric analysis of the measurements which have been made on 136 men, 103 women, and 94 newborn infants, we have proceeded on the conviction that the widest possible usefulness of laboratory investigations of normal human metabolism will result from basing measurements upon those in presumably good health but otherwise typical of human beings in general. It is only when the subjects used for experimentation are representative of the population at large in type, variability and correlation that the results of laboratory research upon limited series may be safely generalized for rationing or for other practical social applications. An explanation of the statistical tests which have been applied to determine the suitability of the series used in the present investigation would lead us into too great detail for this discussion.

The average basal metabolism per twenty-four hours is as follows:

For 136 men .... 1631.74 calories.
For 103 women .... 1349.19 calories.
For 51 male infants .... 144.55 calories.
For 43 female infants .... 140.37 calories.

Thus it appears that the basal energy requirements of the American men are a little less than one half of the number of calories (3,300) established by the Inter-Allied Scientific Food Commission as necessary for rationing in the case of men doing average work eight hours per day. They are a little less than half the 3,200 to 3,600 calories used by a group of men at the Springfield Y. M. C. A. college before they were subjected to rationing tests by the Nutrition Laboratory. The average for new-born infants is somewhat less than ten per cent. of that of women.

Human beings differ in their basal metabolism just as they do in stature, weight, pulse-rate and other measurable characters. For example, the 136 men and 103 women showed the daily caloric output represented by Fig. 1. In these polygons the ordinates represent the frequencies of total heat production in calories per twenty-four hours.

F. 2. FREQUENCIES OF MEN AND WOMEN PRODUCING VARIOUS NUMBERS OF CALORIES PER SQUARE METER OF BODY SURFACE PER TWENTY-FOUR HOURS.

The distribution of daily heat production in men and women is represented by monomodal, more or less symmetrical frequency polygons. This result is of considerable interest since it shows that the distribution of the magnitude of human basal

metabolism exhibits the same orderliness that biometric wor has shown to prevail in the variation of other biological meas urements. When data are somewhat more numerous, it wi be profitable to carry the analysis further by fitting theoretica frequency curves to these frequency distributions.

Measured statistically, the variability of these subjects ma be expressed by standard deviations of 204.66 calories in me and of 155.18 calories per twenty-four hours in women, or b coefficients of variation, *i. e.*, of standard deviations expresse as percentages of the means, of 12.54 per cent. in men and 11.50 per cent. in women.

The statement that variation in the total daily heat produ tion of adults is measured by a coefficient of 11.5 to 12.5 pe cent. will mean very little to the non-statistical reader until h can compare these with coefficients for characters with whic he is more familiar. In our series stature shows a coefficien of variation of 4.39 in men and 3.20 in women, body weight coefficient of variation of 16.06 in men and 20.35 in wome pulse rate at complete rest a coefficient of variation of 10.99 i men and 12.01 in women. Thus basal metabolism shows a var ability far greater than stature but less than body weight an of roughly the same order of magnitude as pulse rate.

Basal metabolism is, therefore, rather highly variable. Th reader will have noted, however, that the foregoing polygon and constants are based upon the total daily heat productions adults in presumably good health but of various body weight statures and ages. It has already been suggested that basa metabolism is related to these physical characters.

We must now inquire whether the observed variability i heat production is due in part to differences in bodily dimer sions. This influence has been considered so great that som physiologists have asserted that heat production per squai meter of body surface is a constant.

That heat production expressed in calories per square met of body surface is not a constant in any exact sense is show by Fig. 2, in which the ordinates represent the frequencies total heat production per square meter of body surface as es mated by the Du Bois height-weight chart.

These diagrams show clearly that heat production p square meter, like total daily heat production, is a variab function. In both cases the frequencies decrease as the magr tudes of the constants diverge more widely, in both the pl and the minus direction, from the average for the whole seri

The fact that a large variability in daily heat producti

remains when it is reduced to calories per square meter of body surface does not, however, warrant the conclusion that metabolism is independent of bodily dimensions.

To investigate this problem we have merely to group all individuals according to some physical character and to determine whether metabolism changes with variation in the magnitude of the selected physical dimension. For example, Fig. 3 is made by representing the stature and the heat production
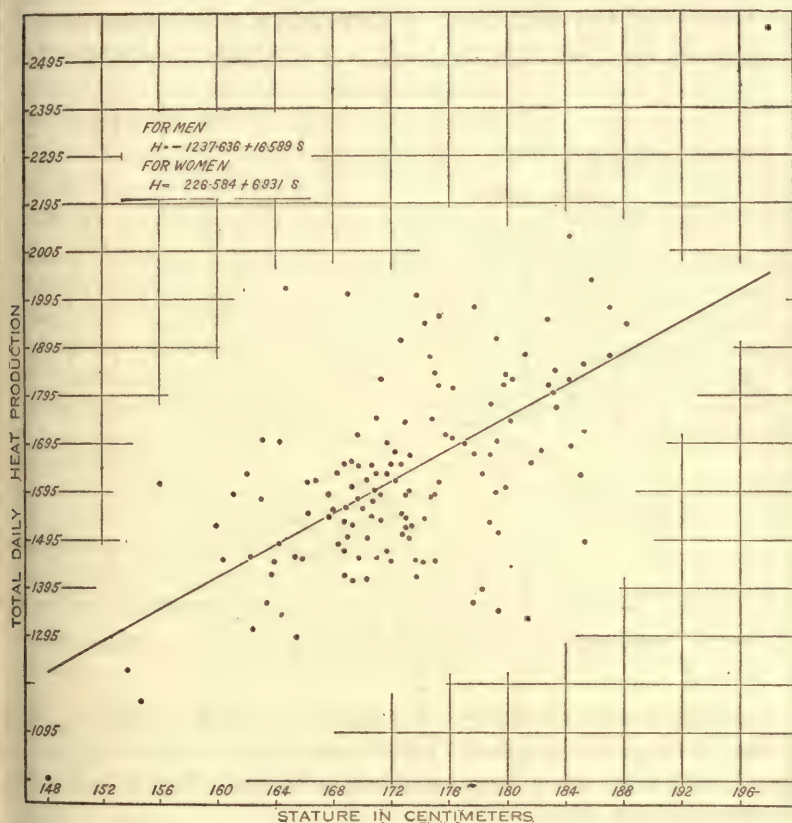

FIG. 3. RELATIONSHIP BETWEEN STATURE AND TOTAL DAILY HEAT PRODUCTION
IN 136 MEN.

of each individual on the horizontal and the vertical scales by a dot. It will be seen at a glance that the metabolism of men of any stature is highly variable. Nevertheless there is an orderly trend in the swarm; there is a marked tendency for taller men to show greater daily heat production.

Such relationships may be represented in a different way. Take another physical character for purposes of illustration. One may determine the average daily heat production of indi-

viduals of different body weights, represent these averages by a series of points and smooth them with a straight line. Thus Fig. 4 shows quite clearly that the daily heat production of individuals tends to increase in a sensibly linear manner with their mass.

For exact comparison we must have recourse to some precise method of expressing the degree of relationship between physical characters and basal metabolism. This may be done by the use of the coefficient of correlation which measures the degree of interdependence of two variables on a universally

FIG. 4. RELATIONSHIP BETWEEN BODY WEIGHT AND DAILY HEAT PRODUCTION.
Actual averages and fitted straight lines.

comparable scale of unity. We find the values for the correlations between stature, body weight and body surface on the one hand, and total daily heat production on the other given in the accompanying table.

Series	Stature and Total Heat-production	Body-weight and Total Heat-production	Body-surface and Total Heat-production
Men (136)	+ 0.615 ± 0.036	+ 0.796 ± 0.021	+ 0.820 ± 0.019
Women (103)	+ 0.232 ± 0.063	+ 0.609 ± 0.042	+ 0.611 ± 0.042
Male infants (51)	+ 0.619 ± 0.058	+ 0.752 ± 0.041	+ 0.749 ± 0.042
Female infants (43)	+ 0.743 ± 0.046	+ 0.808 ± 0.036	+ 0.809 ± 0.036

The coefficients measuring the relationship between body weight and metabolism are in all cases higher than those between stature and metabolism. Body mass is, therefore, a more important factor in determining (in the proximate but not

necessarily in the causal sense) the basal daily heat production of the individual than is a linear bodily dimension such as stature. The correlations between body weight and metabolism and between body surface and metabolism are of approximately the same magnitude. The two characters have, therefore, the same value in indicating the basal metabolism of the subject.

We have just noted that metabolism is correlated with both stature and body weight. Heavier and taller individuals have a larger daily food requirement. But body weight and stature are correlated characters. The interesting question, therefore, naturally arises whether the greater heat production of tall individuals may not be merely the resultant of the relationships between stature and weight on the one hand and weight and metabolism on the other.

This question may be solved by the use of appropriate partial correlation formulæ. We have to determine whether there is a correlation between body weight and daily heat production when correction has been made for the influence of stature, and conversely, to determine whether there is a correlation between stature and daily heat production when correction is made for the influence of body weight. The results are given in the accompanying table.

Series	Correlation Between Weight and Total Heat-production			Correlation Between Stature and Total Heat-production		
	Gross Correlation	Corrected for Stature	Difference	Gross Correlation	Corrected for Weight	Difference
Men (136)	$+0.796 \pm 0.021$	$+0.687 \pm 0.031$	$+0.109$	$+0.615 \pm 0.036$	$+0.321 \pm 0.052$	$+0.294$
Women (103)..	$+0.609 \pm 0.042$	$+0.580 \pm 0.044$	$+0.029$	$+0.232 \pm 0.063$	$+0.045 \pm 0.066$	$+0.187$
Male infants (51)...	$+0.752 \pm 0.041$	$+0.549 \pm 0.066$	$+0.203$	$+0.619 \pm 0.058$	$+0.095 \pm 0.094$	$+0.524$
Female infants (43)...	$+0.808 \pm 0.036$	$+0.494 \pm 0.078$	$+0.314$	$+0.743 \pm 0.046$	$+0.149 \pm 0.101$	$+0.594$

Since the partial correlation coefficients have sensible positive values, it is evident that both stature and body weight have independent significance in indicating daily heat production. This is a result of great importance, since it underlies the determination of the best formulæ for the prediction of the basal metabolism of the individual.

Let us now consider the relationship between metabolism and age.

The change in the food requirements of the human individual with age is not merely a question of material importance to the clinician, but of great interest to the biologist in its bearing upon the general problem of senescence.

The decrease in basal metabolism with age during the period of adult life has been shown to be well represented by the linear equations.

For men ($N=136$),

$$h=1823.80-7.15\,a, \quad h_k=28.703-0.112\,a, \quad h=1022.17-3.60\,a.$$

For women ($N=103$),
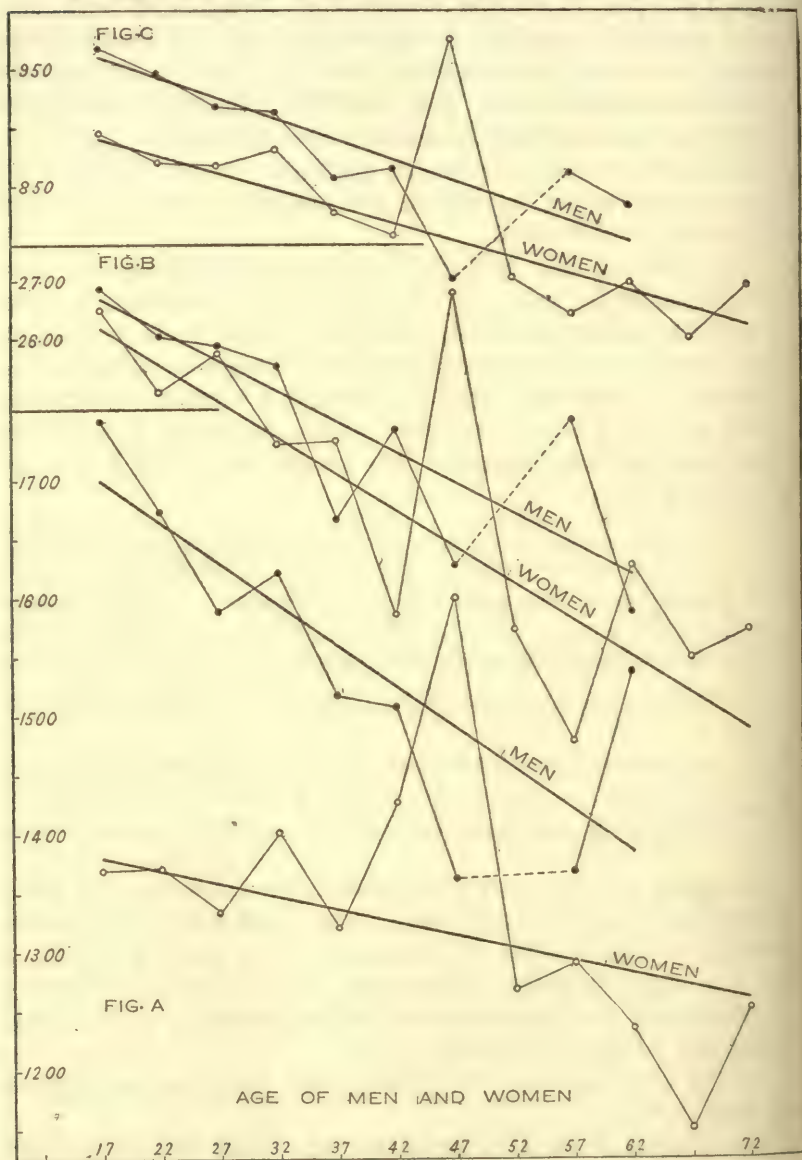
$$h=1420.47-2.29\,a, \quad h_k=28.308-0.124\,a, \quad h=924.25-2.96\,a.$$

FIG. 5. DECREASE IN METABOLISM WITH AGE IN MEN AND WOMEN. Fig. A shows the decrease in total daily heat production, B, the decrease in calories per kilogram of body weight and C, the decrease in calories per square meter of body surface. The straight lines are due to the equations given in the text.

where $h =$ total heat production in calories per twenty-four hours, $h_k =$ calories per kilogram of body weight, $h^D =$ calories per square meter of body surface, as estimated by the Du Bois height-weight chart. These lines are represented in the three diagrams of Fig. 5. While the actual mean heat productions are distributed about these lines with very great irregularity because of the small number of individuals (considered from the statistical, not from the physiological, side) it is doubtful whether equations other than those for straight lines could be advantageously employed.

These equations show that in men the daily heat production decreases about 7.15, while in women it decreases about 2.29 calories per year. Women are smaller than men and have a lower heat production. When the decrease in metabolism with age is expressed in calories per kilogram of body weight or in calories per square meter of body surface, the results for the two sexes are more nearly identical. The linear nature of the change in metabolism with age during the period of adult life fully substantiates the conclusions of biological writers concerning the greater continuity of senescence in the vertebrates as compared with lower organisms. It also shows that during adult life senescent changes take place at a sensibly uniform rate.

Throughout the entire history of the investigation of the metabolism of the warm-blooded animals, the question of the relationship between the body surface area of the organism and its heat production has been a center of interest. Even before the development of adequate experimental methods of investigating the energy transformation which takes place in the body of the vertebrate organism, the possible relationship of body surface area to heat production was a subject of speculation. Physiologists recognized that a whale or an auk in the arctic exposes relatively far less surface for the loss of heat than a flying fish or a humming bird in the tropics. Attempts were consequently made to explain the relatively higher food requirements of small as compared with large animals by the relatively larger surface exposed for heat loss in small organisms.

Newton's law of cooling made a strong appeal to the imagination of earlier physiological writers. It is not surprising that, impressed as they were by the relative constancy of body temperature in the warm-blooded animals, they conceived of heat production as proportional to heat loss as a means of maintaining constant body temperature, and came to look upon heat loss as determined by body surface area, and to consider heat loss as determining in its turn heat production.

This is not the place to trace the history of the so-called body surface law, according to which heat production per square meter of body surface is a constant. It has been widely maintained since the days of the speculative writings of Bergmann, the comparative studies of Rameaux, and the experimental investigations of Müntz, Rubner and Richet. Of recent years this so-called law has assumed practical importance in that writers have maintained that the closest approximation to the basal metabolism of a subject is given by

$$h = \overline{h}_s s,$$

where h is the required daily heat production, $\overline{h}_s$ the average heat production per square meter of body surface in a standard series and s the body surface of the subject under consideration.

The reader may be inclined to inquire why such a formula is of practical significance. The answer should be evident on a moment's reflection. Suppose the clinician wishes to investigate the influence of some disease, for example diabetes, on the metabolism. A subject is selected from the hospital ward, placed in the respiration chamber, and his daily heat production determined. This is merely a technical matter. The interpretation of the result presents a much more serious problem. The caloric output of the subject in a pathological state has no significance as indicating the influence of disease upon metabolism until it can be compared with a normal value. With what normal constant shall it be compared? Naturally, with that which would be expected for the same individual in good health. Thus the *theoretical* metabolism of the subject in health must be taken as a basis of comparison for his actually measured metabolism in disease before one can draw any conclusions whatever concerning the influence of the disease.

Again, suppose that stringent rationing is under consideration. What ration shall be allotted to individuals of various sizes? Clearly, both the most just and the most advantageous procedure would be to allow them food proportional to their physical needs.

Now one of the crucial tests of the validity of a law is its capacity for predicting the unknown. If the "body surface law," expressed by the above formula, serves to predict the heat production of a subject more precisely than any other formula, it must certainly take its place as one of the most important empirical laws in nutritional physiology.

This problem has been investigated in great detail in the extensive data collected at the Nutrition Laboratory during the

past decade. These are now so numerous that it is possible to use one fraction of the records as a basis of prediction equations, and to test the validity of these equations by considering the metabolism of other actually measured individuals as unknown, calculating their caloric output by the various methods, and determining which formula gives the closest prediction. That formula which estimates the metabolism most exactly from measurable physical characters must be looked upon as the most valuable empirical law.

It must be admitted that the "body surface law" has given excellent results. If, however, there be no purely physiological basis for assuming a causal relation between body surface and heat production it would seem desirable, if possible, to replace this formula by a more rational one.

The foregoing analysis has shown that weight, stature and age all have independent significance for predicting the metabolism of the individual. Availing ourselves of the constants showing the independent relationship between these easily ascertainable characters and metabolism, we deduce the following multiple prediction equations:

For men $h = 66.473 + 13.752\,w + 5.003\,s - 6.755\,a.$
For women $h = 665.096 + 9.563\,w + 1.850\,s - 4.676\,a.$

where $w =$ body weight in kilograms, $s =$ stature in centimeters, and $a =$ age in years.

These equations make possible the closest prediction of the daily caloric output of an unknown subject. They are particularly well adapted to practical work. To calculate the most probable metabolism of any subject it is only necessary to substitute the actual values of weight, stature and age in the equation; for example, A. S. F. is a man 21 years old, weighing 69.3 kilograms, and 169 centimeters in height. His most probable daily heat production will therefore be given by

$h = 66.473 + (13.752 \times 69.3) + (5.003 \times 169) - (6.755 \times 21) = 1723$ calories.

His actually measured heat production was 1,733 calories, or there was an error of only 10 calories per twenty-four hours or of 0.6 per cent. in predicting his metabolism from two physical characters and age.

The result is unusually good. K. G. M. is 32 years old, weighs 68.8 kilograms and is 171 centimeters tall. His daily heat production should, therefore, be given by

$h = 66.473 + (13.752 \times 68.8) + (5.003 \times 171) - (6.755 \times 32).$

The equation gives 1,652 calories as compared with 1,889

together. In eleven cases the actual heat productions are higher, while in eleven cases they are lower than the values computed from the equations. The means of the actual and computed heat productions are practically identical. These results show clearly that there is no appreciable influence of vegetarian diet on the basal metabolism.

Other illustrations might be given. Perhaps the most interesting is the use of the equations in investigating the difference in the metabolism of men and women. This problem, which has attracted the interest of a number of investigators in the past, has been reconsidered from all sides on the larger series of data now available. The results show that the average daily (twenty-four hours) basal heat production of the 136 men investigated is 1,632 calories, whereas that of the 103 women studied is 1,349 calories. Thus the daily heat production of women is about 300 calories less than that of men. But women are smaller than men. If correction for body size be made by expressing heat production in calories per kilogram of body weight, it is 25.7 calories in the men as compared with 24.5 calories, or 1.2 calories per kilogram less, in the women. The men show an average daily heat production per square meter of body surface of 925 calories as compared with 850 calories, or 75 calories less, in the women.

The most critical test of the existence of a difference in the metabolism of men and women is that furnished by comparing the actually measured metabolism of women with that calculated from biometric equations on the assumption that they are men of like stature, weight, body surface, age or combinations of these characters. The diagrams in Fig. 8 show the differences between the actual metabolism of women (circles and lower lines) and the heat production calculated on the assumption that they were men of comparable physical characters and age (solid dots and upper lines). Diagrams *A–C* represent the results given by three different equations. The shaded zone shows a deficiency in the actual heat production of the women, who are classified according to body weight, throughout.

These and further statistical tests which can not be discussed in detail show conclusively that the metabolism of American women is lower than that of men. Our results show that the differentiation of the sexes is not evident in infancy. They do not confirm the conclusion of Sondén and Tigerstedt that the difference between men and women tends to disappear with age. Instead we find the difference in the metabolism of men and women well marked throughout the period of adult life.
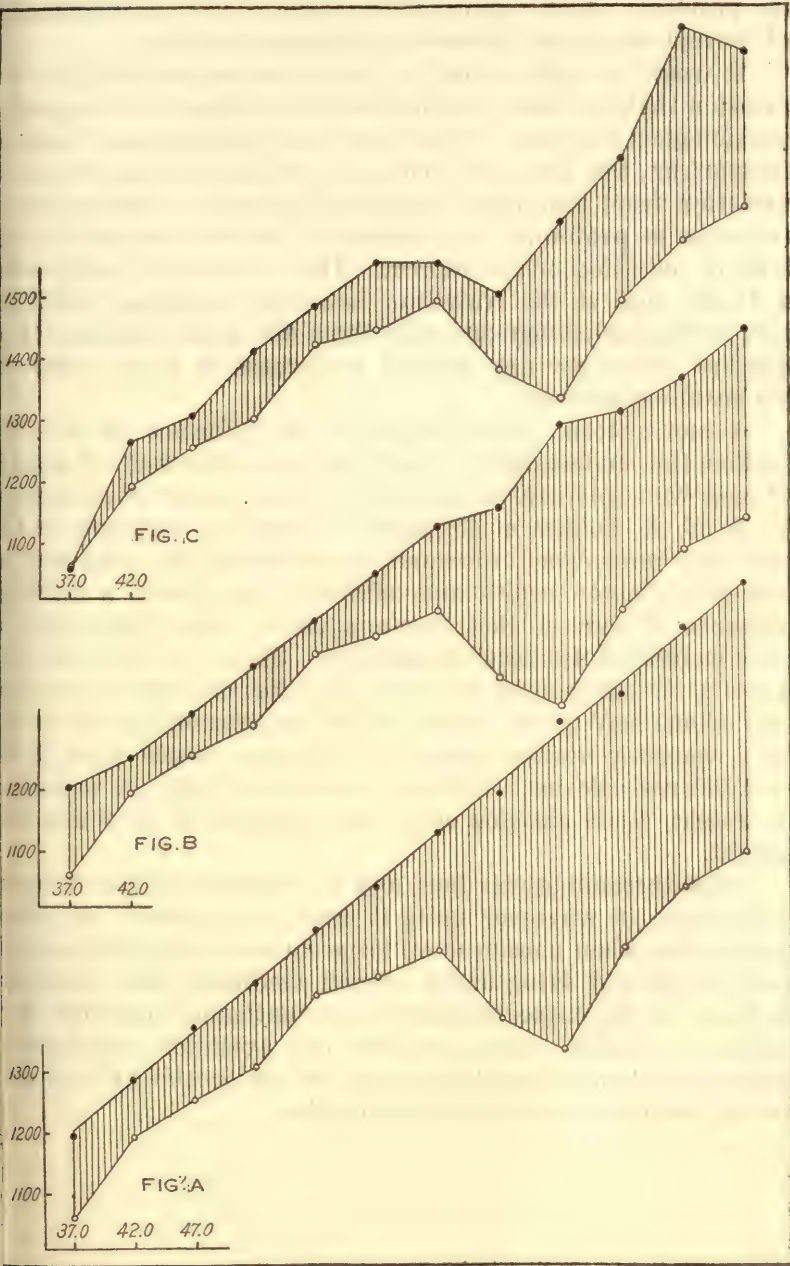
FIG. 8. COMPARISON OF METABOLISM OF MEN AND WOMEN.

This brief outline may serve as an introduction to some of the problems which require consideration in the establishment of normal standards for work in human nutrition.

It would be quite unfair to leave the reader with the impression that the basal metabolism is a fixed and unchangeable physiological constant. While extremely valuable as a basis of comparison, the basal metabolism is subject, not only to great variation from individual to individual, but to modification in response to profound experimentally induced changes in the level of nutrition of the subject. Thus a man who underwent a 31-day fast at the Nutrition Laboratory, during which he took no food whatever and only about 900 cubic centimeters of distilled water per day showed a decrease of 28 per cent. in his basal metabolism.

A more recent investigation of the influence of severely limited diet, undertaken by the Nutrition Laboratory on squads of men who volunteered for this purpose at the International Y. M. C. A. College at Springfield, Mass., in response to the need for more exact information concerning the influence of war-time diet on health and efficiency, has shown a striking influence of reduced diet accompanied by rapid alteration of body weight on the basal metabolism. One squad was kept for a period of four months on a restricted diet with an energy content of one half to two thirds of the requirements prior to the fast, when the normal demand of the men ranged from 3,200 to 3,600 net calories. After a reduction of only 12 per cent. in weight, 1,950 calories only were required to maintain this weight.

Notwithstanding this fact, and the wide variability in basal metabolism in whatever units it may be expressed, the basal metabolism when measured on large numbers of individuals in good health and living under normal conditions, and described in terms of the proper biometric constants and equations, furnishes a valuable, and as yet the only available, standard of comparison in the investigation of all the special problems of energy requirements in human nutrition.

CHARLES BUCKMAN GORING

FEW of the readers of SCIENCE will be familiar with even the name of Charles Goring.[1] His time was largely spent as a prison medical officer. His one monumental work, which may perhaps best be described as *the biology of the convict,* is still unfamiliar to all but a limited circle.

Goring's work[2] was based on thousands of data and is stringently biometric in form, but he was no mere measurer, card shuffler and constant computer. He knew his convicts as the trained student of animal behavior knows his organisms—and better, for he had not merely their physical measurements and an intimate personal knowledge and evaluation of their mental characteristics but knew much of their ancestry and family associations. To Goring, measurements were inviolate—not to be juggled with, modified or discarded because they did not substantiate a popular theory.

[1] Goring was born in 1870 and died in 1919. He was a student and later a fellow of University College, London. He served on a hospital ship during the Boer War. At the time of his death—met at his post combating the influenza epidemic—he was Medical Officer in Chief at Strangeways Prison, Manchester. Those who desire may find a portrait and a more adequate appreciation in *Biometrika,* Vol. XII., pp. 297–307, pl. 1, 1919.

[2] Goring, C. B., "The English Convict; A Statistical Study." 444 pp. London, 1913. Abridged edition, Wyman and Co., 1915. The statistical work on this volume was carried out at the Biometric Laboratory with the cooperation of H. E. Soper and with the helpful suggestion and criticism of Professor Pearson.

Better proof of this could not be found than the fact that the raw data for his book were set up before the calculations were well under way. Goring as a thoroughgoing biometrician believed that in many fields of research valid conclusions must rest upon the mathematical analysis of large masses of data. But in his research each constant was critically weighed against his own broad and intimate personal experience of the individual instances which constitute the mass.

I find it difficult to decide just what characteristic of Goring impressed me most when we were working together at the Biometric Laboratory ten years ago. Sometimes it was the steadfast scientific purpose which had supported the years of painstaking detail upon which his great book rests—detail scrupulously executed notwithstanding the fact that there was at times little prospect of its ever serving as a basis for constants and generalizations. Sometimes it was the breadth of interests, knowledge and sympathies of one whose work lay in a field seemingly so circumscribed. Sometimes it was the entire freedom from both callousness and sentimentality of a man who had spent a decade, more or less, with the inmates of the British prisons.

One sentence tells much of the man. One day I asked, " Why is this to be *The English Convict* instead of *The English Criminal?* " He replied instantly, " Perhaps some of them are not criminals, only convicts."

<div style="text-align: right">J. Arthur Harris</div>

THE VARIATION AND THE STATISTICAL CONSTANTS OF BASAL METABOLISM IN MEN

BY

J. ARTHUR HARRIS AND F. G. BENEDICT

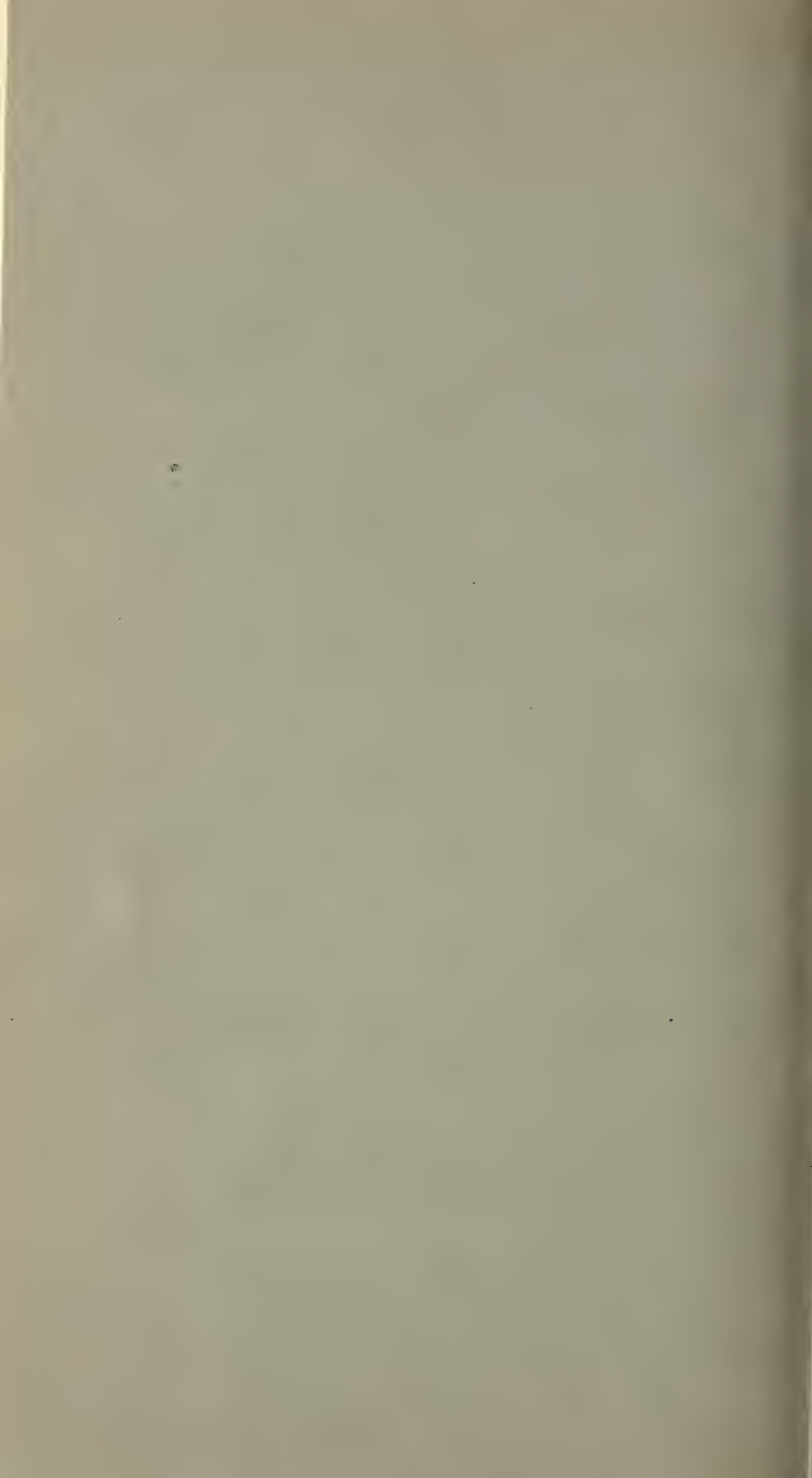# THE VARIATION AND THE STATISTICAL CONSTANTS OF BASAL METABOLISM IN MEN.

By J. ARTHUR HARRIS and F. G. BENEDICT.

(*From the Nutrition Laboratory and the Station for Experimental Evolution, the Carnegie Institution of Washington.*)

(Received for publication, January 6, 1921.)

I. INTRODUCTION.

Physiologists now generally accept the so called basal metabolism of the individual as a standard in all comparisons involving the consideration of energy transformation. Furthermore, the technical conditions under which measurements of the basal metabolism shall be made are now generally agreed upon. The opinion has even been widely maintained that heat production per square meter of body surface area is a constant.

A detailed review and discussion of the literature and the results of a biometric analysis of a series of basal metabolism data for 136 men and 103 women have been given elsewhere.[1] The results of this analysis show the following coefficients of variation for the averages[2] of the daily basal metabolism constants for the individual subjects.

	Men.	Women.
Total calories per 24 hrs.	12.54	11.50
Calories per kg. per 24 hrs.	9.36	14.14
Calories per square meter per 24 hrs.	8.05	9.17

[1] Harris, J. A., and Benedict, F. G., *Carnegie Inst. of Washington, Pub.* 79, 1919, 129–200.

[2] These are averages of daily means in all cases in which measurements could be made on more than 1 day. For 35 subjects a single daily mean only was available.

In practise the catabolism of a subject is almost invariably measured in two or more short periods on a given day in a respiration apparatus. Technically these are generally designated as *minimum periods*, since the lowest values are generally found in the absence of muscular activity. The term somewhat misleading when averages of two or more periods are taken,

These results indicate a rather wide variability in basal metabo
lism from individual to individual, even when heat production i
corrected for body size by expressing it in calories per kilo o
body weight or in calories per square meter of body surface area

The problem now arises whether, under conditions such tha
the age factor[4] is practically eliminated, the basal metabolism c
the individual is essentially constant from day to day or whether i
shows sensible variations. In 1915 one of us[5] discussed the ques
tion of changes in the metabolism of the individual from day t
day, and showed that, contrary to the early assumption that th
basal metabolism of an individual remains essentially constan
from day to day, it is really variable.[6]

since there can be only one term minimum. The calories produced or th
volume of the gaseous exchange for any one of these periods which seems t
those acquainted with the technical details of the experiment reasonabl
free from experimental error, might serve as a measure of the metabolism
Since, however, the basal metabolism of the individual can probably b
assumed to be sensibly constant for 1 day, it seems wisest to regard th
individual periods which appear the most free from possible experiment
error as of the nature of duplicate, triplicate, etc. analyses, and to averag
these values. This daily mean has served as the unit in work at the Nutr
tion Laboratory.

[3] The frequency distributions of total calories and calories per squar
meter are shown in Figs. 1 and 2 of Harris and Benedict (Harris, J. A., an
Benedict, F. G., *Scient. Month.*, 1919, viii, 388–389).

[4] The results of a previous study of the relationship between age an
metabolism (Harris and Benedict,[1] pp. 107–127) have shown that there
a gradual and practically linear decrease in metabolism with age durin
the period of adult life. The annual rate of decrease amounts to about 7.1
calories (per 24 hrs.) per year in men and 2.29 calories (per 24 hrs.) per yea
in women. Correction for body size by expressing the results in calori
per kilo of weight gives a rate of decrease of 0.112 calorie in men and 0.1
calorie in women. If the results are expressed in calories per square met
of body surface area as estimated by the Du Bois height-weight chart, t
decrease for men is 3.60 calories and for women 2.96 calories per 24 hrs.
life. The change due to age is not, therefore, large and cannot be assum
to be an important factor for subjects observed during a short period
time only.

[5] Benedict, F. G., *J. Biol. Chem.*, 1915, xx, 290–295.

[6] This conclusion was based on a study of the difference between t
highest and the lowest daily basal metabolism constant expressed as
percentage of the minimum value as a base. The study showed that t
differences varied greatly from subject to subject. In one case the ma
mum and minimum oxygen consumption varied as widely as 31.3 per ce

The purpose of the present paper is threefold:

1. To obtain some measure of the variability of the basal metabolism of the normal individual.

2. To consider the relationship between the length of time over which the observations extend and the variation in the metabolism of the individual.

3. To consider the most suitable method for determining the population mean from measurements on a series of individuals.

II. PRESENTATION OF RESULTS.

1. Variation of Metabolism in the Individual.

For a certain number of men investigated at the Nutrition Laboratory the number of days on which measurements were made is sufficiently large (20 or more) to justify the calculation of statistical constants for the individual subjects.

These appear in Tables I to III.

The constants indicate that individuals are differentiated among themselves with respect to basal metabolism even when expressed in calories per kilo of body weight or calories per square meter of body surface area as estimated by the Du Bois height-weight chart.[7]

Furthermore it is clear that the metabolism of each individual subject is to a considerable degree variable. This is shown by the rather wide range between the maximum and minimum daily metabolism for each individual as shown in these tables. These ranges are expressed as percentages of the minimum value found in the tables. It is also evident in the absolute variabilities as

the minimum value. In another case it varied only 3.5 per cent. While it was pointed out that it was hardly correct to obtain an average value for the oxygen consumption for individuals with such wide differences in the time covered by the observations, an average value was determined in the absence of any better available method, and found to be 13.9 per cent.

[7] Since a discussion of the differentiation of individuals with respect to basal metabolism is not a primary purpose of this paper the subject is not pursued farther. The statement above may be verified by taking differences between the various constants and comparing them with their probable errors. It is to be noted that these constants are uncorrected for age, and that the age differences will, in general, tend to increase slightly the differentiation of the subjects.

given in terms of the standard deviations (S. D.). Expressin
the total amount of variation as measured in terms of the stand
ard deviation as a percentage of the means we have the relativ
variabilities expressed as coefficients of variation $\left(\text{C.V.} = \dfrac{100\,\text{S.D.}}{\text{Mean}} \right.$

TABLE I.

Statistical Constants for Basal Metabolism in Eleven Men.

No. and individual.	Days.	Total calories per 24 hrs.					Coefficient of variation
		Minimum.	Maximum.	Percentage range.	Mean.	Standard deviation.	
47. F. P. R....	20	1,446	1,684	16.5	$1{,}540.0 \pm 11.2$	74.6 ± 8.0	4.
45. K. H. A....	25	1,505	1,765	17.3	$1{,}648.3 \pm 9.2$	68.2 ± 6.5	4.
96. A. J. O....	25	1,679	1,804	7.4	$1{,}741.5 \pm 4.5$	33.5 ± 3.2	1.
41. C. B. S....	26	1,592	1,785	12.1	$1{,}699.2 \pm 6.5$	49.0 ± 4.6	2.
61. J. K. M....	27	1,458	1,650	13.2	$1{,}546.6 \pm 6.6$	51.2 ± 4.7	3.
54. H. H. A....	28	1,327	1,686	27.1	$1{,}488.3 \pm 10.3$	80.7 ± 7.3	5.
66. L. E. E....	31	1,596	1,848	15.8	$1{,}705.6 \pm 7.9$	64.9 ± 5.6	3.
59. H. L. H. ..	35	1,548	1,890	22.1	$1{,}694.4 \pm 9.2$	80.4 ± 6.5	4.
70. H. F. T....	41	1,205	1,514	25.6	$1{,}350.4 \pm 7.7$	73.1 ± 5.4	5.
9. M. A. M...	53	1,562	1,917	22.7	$1{,}696.0 \pm 6.9$	74.7 ± 4.9	4.
48. J. J. C....	53	1,511	1,740	15.2	$1{,}583.8 \pm 4.4$	47.3 ± 3.1	3.

TABLE II.

Statistical Constants for Basal Metabolism in Eleven Men.

No. and individual.	Days.	Calories per kg.					Coefficient of variation
		Minimum.	Maximum.	Percentage range.	Mean.	Standard deviation.	
47. F. P. R....	20	22.1	25.7	16.3	23.65 ± 0.16	1.08 ± 0.12	4.
45. K. H. A....	25	22.8	27.1	18.4	24.84 ± 0.15	1.10 ± 0.10	4.
96. A. J. O....	25	23.3	26.6	14.2	25.13 ± 0.11	0.81 ± 0.08	3.
41. C. B. S....	26	21.9	25.4	16.0	23.92 ± 0.11	0.81 ± 0.08	3.
61. J. K. M....	27	24.5	27.4	11.8	25.63 ± 0.09	0.70 ± 0.06	2.
54. H. H. A....	28	22.2	26.9	21.2	23.89 ± 0.14	1.10 ± 0.10	4.
66. L. E. E....	31	26.8	31.5	17.5	28.47 ± 0.14	1.15 ± 0.10	4.
59. H. L. H....	35	25.8	31.5	22.1	28.05 ± 0.16	1.39 ± 0.11	4.
70. H. F. T....	41	21.1	26.0	23.2	23.35 ± 0.12	1.18 ± 0.09	5.
9. M. A. M...	53	23.8	28.5	19.7	25.72 ± 0.10	1.04 ± 0.07	4.
48. J. J. C....	53	22.6	26.7	18.1	24.39 ± 0.07	0.81 ± 0.05	3.

We note that the coefficients for total calories range from 1.9 to 5.4 per cent with a general average of 3.97. Those for calories per kilo of body weight range from 2.8 to 5.1 per cent with a general average of 4.40, and those for calories per square meter of body surface from 2.3 to 5.3 per cent with a general average of 3.95.

The suggestion will naturally arise that the variation in basal metabolism in these longer periods is due merely to the uniform decline in metabolic activity characteristic of adult life.

TABLE III.

Statistical Constants for Basal Metabolism in Eleven Men.

No. and individual.	Days.	Calories per square meter.					
		Minimum.	Maximum.	Percentage range.	Mean.	Standard deviation.	Coefficient of variation.
F. P. R....	20	808	936	15.8	864.0±6.1	40.8±4.3	4.7
K. H. A....	25	814	954	17.2	885.6±5.0	36.9±3.5	4.2
A. J. O....	25	879	970	10.4	927.2±2.9	21.5±2.0	2.3
C. B. S....	26	834	947	13.5	898.2±3.6	27.6±2.6	3.1
J. K. M....	27	851	948	11.4	897.4±3.5	26.7±2.5	3.0
H. H. A....	28	804	998	24.1	884.3±5.6	43.9±4.0	5.0
L. E. E....	31	917	1,074	17.1	980.5±4.7	38.6±3.3	3.9
H. L. H....	35	905	1,105	22.1	986.5±5.4	46.9±3.8	4.8
H. F. T....	41	700	870	24.3	779.2±4.3	40.9±3.0	5.3
M. A. M...	53	863	1,042	20.7	935.2±3.6	39.0±2.6	4.2
J. J. C.....	53	840	956	13.8	883.8±2.4	26.3±1.7	3.0

Evidence that this is not the case will be adduced in the following section in which it will be shown that for periods of not more than 10 or 15 days the magnitude of the variation in metabolism is positively correlated with the duration of the period over which the observations extended. We have, furthermore, applied a correction for age[8] to the measurements of three of the individuals with results as given in Table IV. The standard deviations and the coefficients of variation are practically the same after correction for age as for the original observations. The result shows clearly that age is not a primary factor underlying the variations.

[8] Calculations from unpublished data.

TABLE IV.

Comparison of Basal Metabolism Constants Corrected and Uncorrected j Age.

Subject and range.	Constant uncorrected for age.	Constant corrected for age.	Differen in constan
88. T. M. C., range = 1,624.			
Standard deviation....................	54.04	53.27	−0.7
Coefficient of variation................	4.18	4.12	−0.0
59. H. L. H., range = 905.			
Standard deviation....................	80.39	82.07	+1.6
Coefficient of variation................	4.74	4.84	+0.1
48. J. J. C., range = 819.			
Standard deviation....................	47.27	47.15	−0.1
Coefficient of variation................	2.98	2.98	±0.0

2. The Time Factor in the Variation of the Metabolism of the Individual.

In the preceding section we dealt with the problem of the var- bility within the individual on the basis of data for a few m upon whom more extensive series of measurements had be made. By the application of other methods it is possible to pu the analysis somewhat farther.

While the ultimate purpose of studies of variation in the met: olism constants of the same individual should be to determ something of the proximate causes underlying these variatio, it is worth while to obtain some general idea of the amount f variation which may be expected to occur in the individual subj t with a lapse of time.[9]

---

[9] In the first discussion of this subject (Benedict,[5] p. 292) the sim e method of range of variation in metabolism led to the following conclusi : "A general inspection of the data will show that, as a rule, the great t variations were found with the subjects studied over the longest peric. While it is hardly correct to obtain an average value for the oxygen consui - tion for so many different individuals with such wide differences in the t e covered by the experiments, yet such a value has been found and shows t t on the basis of these observations there may be an average variatior f 13.9 per cent in the basal metabolism, when measured in the post-absorpi e condition and with complete muscular repose, during a period of two ye s or, in the majority of cases, considerably less. With no attempt to anal e the causes of these differences, it is sufficient here simply to call attent n to their magnitude."

If the metabolism of the individual changes from time to time irrespective of changes in bodily dimensions it would seem reasonable to presume that these changes would be greater for more widely separated periods. Thus the observations made on a single day can be reasonably regarded as based upon a subject in practically stationary physical and physiological conditions. Those made at widely separated dates more probably represent the individual in somewhat different physiological states. Thus while the active protoplasmic mass is probably essentially identical in the two cases (if the periods are not too widely separated) the unknown stimulus to metabolic activity may differ to a considerable extent from one period to the other.

As a measure of variation of the metabolism of the individual we have adopted the standard deviation[10] of the measurements of each subject. As a measure of time covered by the observation we have taken the actual number of days, including the days upon which the measurements were made, i.e. $(T_2 + 1) - T_1$, where T_1 and T_2 are the times of the first and last measurements. Thus if the metabolism of an individual were measured on July 1 and 2 the range would be 2 days. If three observations were made, one on July 1, one on July 10, and one on September 3 of the same year, the range would be 65 days. Correlating between the range in days and the standard deviation of total calories per 24 hours we have for 101 individuals

$$r = +0.276 \pm 0.062, \quad \frac{r}{E_r} = 4.45$$

The coefficient measuring the relationship between range in days and the standard deviation of calories per square meter as estimated by the Du Bois height-weight chart, which will be the only approximation to the body surface used, is

$$r = +0.254 \pm 0.063, \quad \frac{r}{E_r} = 4.03$$

The correction between the range of days over which the experiments extended and the standard deviation of calories per kilo of body weight is

$$r = +0.248 \pm 0.063, \quad \frac{r}{E_r} = 3.94$$

---

[10] The coefficient of variation might have been used with equal propriety.

TABLE V.

Correlation between Duration of Period of Observation and the Stand *Deviation of Basal Metabolism in Calories* **per** *Kg. per 24 Hrs.*

Range.	No. of subjects.	Correlation.		Mean standard deviation for whole range.	Mean standar deviation individu added
		$r \pm E_r$	$\dfrac{r}{E_r}$		
days					
2–5	19	$+0.509 \pm 0.115$	4.43	0.463	
2–10	32	$+0.398 \pm 0.100$	3.98	0.544	0.66
2–15	37	$+0.297 \pm 0.101$	2.94	0.555	0.62
2–20	42	$+0.343 \pm 0.092$	3.73	0.591	0.85
2–30	49	$+0.232 \pm 0.091$	2.55	0.599	0.64
2–50	59	$+0.139 \pm 0.086$	1.62	0.605	0.63
2–125	73	$+0.266 \pm 0.073$	3.64	0.648	0.82
2–160	81	$+0.344 \pm 0.066$	5.21	0.679	0.96
2–240	86	$+0.363 \pm 0.063$	5.76	0.697	0.97
2–360	90	$+0.370 \pm 0.061$	6.07	0.709	0.97
2–370	91	$+0.327 \pm 0.063$	5.19	0.708	0.58
2–380	92	$+0.337 \pm 0.062$	5.44	0.712	1.05
2–415	93	$+0.349 \pm 0.061$	5.72	0.716	1.09
2–425	94	$+0.334 \pm 0.062$	5.39	0.717	0.81
2–430	95	$+0.314 \pm 0.062$	5.06	0.717	0.70
2–600	96	$+0.280 \pm 0.063$	4.44	0.717	0.70
2–775	97	$+0.297 \pm 0.063$	4.71	0.721	1.14
2–800	98	$+0.245 \pm 0.064$	3.83	0.720	0.57
2–820	99	$+0.231 \pm 0.064$	3.61	0.720	0.80
2–905	100	$+0.280 \pm 0.062$	4.52	0.727	1.39
2–1,624	101	$+0.248 \pm 0.063$	3.94	0.729	0.91

TABLE VI.

Correlation between Duration of Period of Observation and the Stana *Deviation of Basal Metabolism in Calories per Kg. per 24 Hrs.*

Range.	No. of subjects.	Correlation.	
		$r \pm E_r.$	$\dfrac{r}{E_r}$
days			
332–1,624	12	$+0.116 \pm 0.192$	0.60
236–1,624	16	$+0.018 \pm 0.169$	0.11
159–1,624	21	-0.067 ± 0.147	0.46
131–1,624	28	-0.014 ± 0.127	0.11
121–1,624	31	$+0.007 \pm 0.121$	0.06
85–1,624	37	$+0.020 \pm 0.111$	0.18
51–1,624	42	$+0.062 \pm 0.104$	0.60

5-1,624 days) are positive in sign, as are two of the other four coefficients. There is, therefore, a suggestion of positive correlation in the longer time groups. It is clear from the magnitude of these correlation coefficients that, as suggested above, the standard deviations do not steadily increase with the lapse of
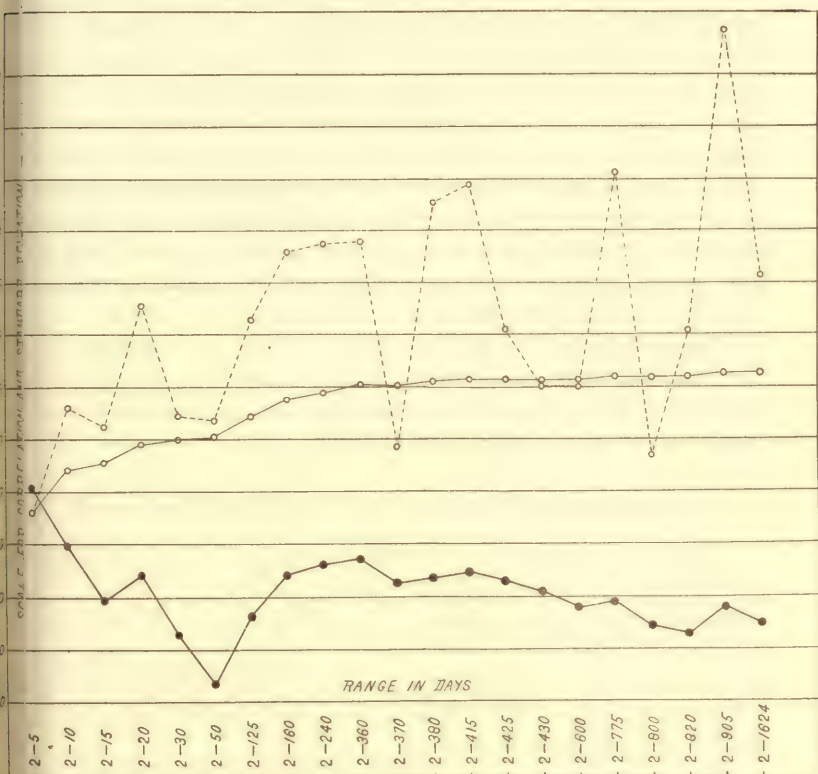

Fig. 1. Abscissæ represent maximum range in days; ordinates show standard deviations of daily means in the individual or correlations between range of time covered by observations and standard deviation of daily metabolism. Solid dots and line represent correlation coefficients; circles and solid line represent the mean standard deviations for the individuals of each group for which the correlations were computed; circles and broken line represent the mean standard deviation of the individuals added to each preceding group to form the group upon which the correlation coefficient is based. Because of the great variation in the length of the periods no attempt has been made to represent abscissæ on a uniform scale.

time.   This may be shown to be the case by determining th
average values of the standard deviations of the metabolism o
the groups of individuals for whom the correlations in Table
were calculated.   These means have been calculated for th
groups of 19 to 101 subjects with the range indicated in Table
and for the individuals added to each group to give the next wide
range group.   These mean standard deviations are shown in th
last two columns of Table V.   The first series is represented i
Fig. 1 by the circles connected by continuous lines while the sec
ond is shown by the circles connected by broken lines.   Th
wide variability of the latter averages is attributable to the ver
small number of individuals added in the longer range periods.[14]
Both series of means show that the standard deviation at firs
increases rapidly as the length of the period becomes longer bu
the rate of increase becomes smaller as the periods over whicl
the observations extend increase in length.

The last paragraph on page 264, written as it stands befor
the calculation of the constants in this table, is therefore full
substantiated by the available data for individuals who have bee
studied over longer periods of time.

3. The Statistical Constants for Basal Metabolism in Man.

A primary requisite for continual progress in the investigatio
of human metabolism under various special conditions, e.g. o
exercise, rationing, or disease, is the determination of some stanc
ard which may serve as a basis of comparison.[15]   Such a cor
stant, based upon a considerable series of individuals and designe
to serve as a standard of comparison, may be conveniently desig
nated as a *population constant.*

The average value of the basal metabolism constant has gene
ally served this purpose, whether expressed in actual calorie
per unit of time uncorrected for the influence of body mass, i
calories per kilo of body weight, or in calories per square mete

[14] In the twelve groups in which the range is wider than 2–360 days tl
positions of the circles represent the standard deviations for single ind
viduals only, not means.

[15] The general principles underlying the establishment of control seri
have been discussed elsewhere (Harris and Benedict,[1] pp. 223–227; als
Harris and Benedict,[3] p. 385).

f body surface area as estimated by some one of the various
ormulas which have been proposed for this purpose.

When observations for a number of days are available, and
vhen it is desirable to define the basal metabolism of the indi-
idual more precisely than can be done as the result of experimen-
ation for 1 day, the question naturally arises whether the *aver-
ge* of the *daily means* or the *minimum daily mean* shall be used
o represent the basal metabolism of the individual. The exact
nethod of calculating the average and the standard deviation
herefore requires consideration.[16]

In a preceding publication[1] we used the average of the daily
neans for the determination of the population constant. Our
rotocols of data[17] show these values and indicate the number of
lays and the total number of periods upon which they were based.

One of our helpfully critical correspondents has suggested that
t may be quite improper to lump together and treat as of equal
alue basal metabolism constants for individual subjects some of
vhich rest upon an observation for only a single day while others
epend upon many days measurements.

This criticism has probably also occurred to others. It seems
esirable, therefore, to consider the problem of the best method
f deducing a constant for a series of individuals—a population
onstant—from the experimental readings.

There are seven possible ways of determining the statistical
onstants of a series of metabolism measurements some of which
st upon work for but a single day while others depend upon
veral days observation.

A. The average of the daily averages may serve as the units
presenting the individuals.

If the measured metabolism is variable it seems illogical, as our
rrespondent suggests, to treat the constant obtained from an
servation for a single day as equal in value to that deduced from
number of days observation. The selected constant may, there-
re, be (1) used only once in the calculation of the statistical
nstant for the population; (2) may be weighted with the num-

---

[16] While we have elsewhere proposed the use of multiple prediction equa-
ons for calculating the basal energy requirements, the consideration of
e most suitable method for the determination of the statistical constants
represent the individual is pertinent, since the calculation of these equa-
ons involves the determination of the means and standard deviations.

[17] Harris and Benedict,[1] Tables A to D.

ber of days observation; or, (3) weighted with the square root o
the number of days observation.

B. The minimum value of the daily averages, *i.e.* the constan
for the single day giving the lowest average, may be used.

This method has for its justification a physiological consideration
By definition the basal metabolism is the catabolism in the ab
sence of muscular activity and the stimulatory influence of re
cently ingested food. Since these are the most potent factors i
determining the superbasal metabolism of the individual, th
basal metabolism is, practically speaking, synonymous with th
minimum metabolism. It is possible, therefore, to consider tha
the absolute minimum for any individual should be taken as th
true basal value. If the mean of all the daily averages of meas
urements made upon an individual is used as a measure of th
basal metabolism a value somewhat higher than the absolut
minimum is obtained and we admit that the individual may fa
below his own basal value. Against this method is to b
urged the criticism that the lowest value may be really subbasa
because of errors of measurement. This criticism is in large pal
met by the concordance of the results for the two or more perioc
upon which the daily averages are almost invariably based
These minimum values may, like the individual means, be (4
used only once in the calculation of the statistical constant fo
the population; (5) may be weighted with the number of day
observation; or (6) weighted with the square root of the numbe
of days observation.

C. (7) The constants may be computed directly from the who
series of daily means available.

The first group (1 to 3) may be conveniently designated as th
method of individual means, the second group as the *method of i
dividual minima,* and the third group as the *method of daily mean*

Table VII gives the statistical constants for the daily (24 hour
heat production of the 136 men[18] for whom data (individual mean
are given in the protocols of our former publication.[19]

[18] The 103 women considered in our volume were not studied over
period sufficiently long to make it worth while to calculate weighted co
stants comparable with those for men.

[19] In the full revision of the data for the 863 individual periods a f
minor inaccuracies, of no practical importance for the purposes of o
earlier volume, were found in the fundamental protocols. The unweight
constants have, therefore, been recalculated for the purposes of tl
paper.

271

	Mean.	Standard deviation.	Coefficient of variation.	Mean.	Standard deviation.	Coefficient of variation.	Mean.	Standard deviation.	Coefficient of variation.
1. Unweighted*	1,630.98±11.86	205.01± 8.38	12.57	25.691±0.139	2.402±0.098	9.35	925.147±3.877	67.038±2.742	7.25
2. Weighted with the square root of the no. of days	1,617.44±10.78	186.43± 7.62	11.53	25.766±0.128	2.219±0.091	8.61	921.715±3.753	64.891±2.654	7.04
Difference (2)–(1)	−13.54±16.03	−18.58±11.33	−1.04	+0.075±0.189	−0.183±0.134	−0.74	−3.432±5.396	−2.147±3.816	−0.21
Diff./E diff.	0.84	1.64		0.39	1.37		0.64	0.56	
Percentage difference	0.83	9.49		0.29	7.94		0.37	3.25	
3. Weighted with the no. of days	1,602.60± 9.61	166.22± 6.80	10.37	25.779±0.117	2.030±0.083	7.87	916.272±3.655	63.190±2.584	6.90
Difference (3)–(1)	−28.38±15.26	−38.79±10.79	−2.20	+0.088±0.182	−0.372±0.128	−1.48	−8.875±5.328	−3.848±3.768	−0.35
Diff./E diff.	1.86	3.60		0.48	2.91		1.67	1.02	
Percentage difference	1.76	20.90		0.34	16.81		0.96	5.91	
4. Based on daily observation	1,602.57±10.16	175.61± 7.18	10.96	25.780±0.129	2.227±0.091	8.64	916.282±4.100	70.888±2.899	7.74
Difference (4)–(1)	−28.41±15.62	−29.40±11.04	−1.61	+0.089±0.189	−0.175±0.134	−0.71	−8.865±5.643	+3.850±3.990	+0.49
Diff./E diff.	1.82	2.66		0.47	1.31		1.57	0.96	
Percentage difference	1.76	15.45		0.34	0.76		0.96	5.58	

* For (1), N = 136; for (2), N = 289.4; for (3), N = 863; for (4), N = 863.

Considering first the mean total daily heat production, which is the fundamental constant for the establishment of a standard value, we note that the mean obtained by weighting is somewhat lower than that secured by giving each individual equal weight irrespective of the number of days on which observations were made.

The differences are, however, of a low order of magnitude as compared with the average heat production of 1,631 calories. The heat production is on the average 13.54 calories lower when the constants are weighted with the square root of the number of days and 28.38 calories lower when the constants are weighted with the number of days, or based upon the constants for the individual days.[20]

These differences are between total daily heat productions (unweighted) of about 1,631 calories. Thus they are relatively small only 0.83 and 1.76 per cent[21] by the two methods of weighting. The differences are not merely relatively small as compared with the total heat production but are in all cases less than twice as large as the probable errors of the differences.[22]

Turning now to the results for heat production per kilo of body weight, we have the comparisons set forth in the second section of Table VII.

The means show a slight but wholly insignificant increase in calories per kilo as a result of weighting with the square root of the number of days, or with the number of days, or by using the daily averages in calculating the constants.

Finally consider the results for calories per 24 hours per estimated square meter of body surface.

The heat production is 3.43 calories lower when weighted by the square root of the number of days observation and 8.88 calories lower when weighted with the number of days, than when calculated from the daily averages. The differences are rela-

[20] The means calculated in these two ways should be identical. The slight difference is due to the number of significant figures retained in the calculations.

[21] Percentage differences have been computed by using the average of the two means compared as a base.

[22] The probable errors have in all cases been based on the actual, not the weighted, number of individuals as N.

vely small, being less than 1 per cent in the three comparisons.
ll differences are less than twice as large as their probable errors.

It is clear from the foregoing constants that *practically it is
immaterial whether the population means are calculated from the aver-
ges of the individual subjects, from the averages weighted with the
number of days, or with the square root of the number of days, or
whether they are determined directly from the daily observations.*[23]

From Table VII it appears that the standard deviations obtained
y weighting the individual means with the square root of the
number of days or with the number of days are lower than those
alculated without weighting.

We now have to consider the constants deduced from the mini-
mum values of the daily metabolism. The results are given in
Table VIII. Limiting our attention for the moment to a com-
parison of these constants among themselves we note that, in
whatever units measured, the mean metabolism calculated by
weighting with the square root of the number of days is always
wer than the constant obtained without weighting. When the
minima for the individuals are weighted with the number of
ays instead of with the square root of the number of days the
ifference between the weighted and the unweighted value is even
reater, amounting to -69.6 calories of total daily heat produc-
on, -0.579 calories per kilo, and -32.03 calories per square
meter of surface area. These differences correspond to relative
ifferences of 4.49, 2.35, and 3.63 per cent of the unweighted con-
stants. *They show that if an absolute minimum, i.e. the one single
ay with the lowest average of metabolism measurements, for each
individual is adopted, the constants for a population will depend to
considerable extent upon the number of days observation for each
individual.*

Table IX compares the percentage change in the population
onstant due to weighting when the population constant is calcu-
ated from means and from minima. For all three units of metab-

---

[23] While this is the result for the large series of data in hand the calcu-
ation of the population constant from daily observations by weighting
ith the number of days is not to be generally recommended since in series
in which the number of individuals is small the population average may be
so greatly influenced by repeated observations on one or a few intensively
studied individuals.

TABLE X.

Comparison of Statistical Constants for Basal Metabolism Based on Individual Means and Individual Minima for 196 Men.

	Unweighted. N = 136			Weighted with the square root of the no. of days. N = 289.4			Weighted with the no. of days. N = 863		
	Mean.	Standard deviation.	Coefficient of variation.	Mean.	Standard deviation.	Coefficient of variation.	Mean.	Standard deviation.	Coefficient of variation.
Total calories.									
From means....	1,630.98±11.86	205.01± 8.38	12.57	1,617.44±10.78	186.43± 7.62	11.53	1,602.60± 9.61	166.22±6.80	10.37
" minima....	1,585.11±11.90	205.65± 8.41	12.97	1,550.86±11.01	190.38± 7.79	12.28	1,515.49± 9.91	171.29±7.01	11.30
Difference........	+45.87±16.80	−0.64±11.87	−0.40	+66.58±15.41	−3.95±10.90	−0.75	+87.11±13.80	−5.07±9.77	−0.93
Diff./E_{diff}....	2.73	0.05		4.32	0.36		6.31	0.52	
Percentage difference........	2.85	0.31		4.20	2.10		5.59	3.00	
Calories per kg.									
From means....	25.691±0.139	2.402±0.098	9.35	25.766±0.128	2.219±0.091	8.61	25.779±0.117	2.030±0.083	7.87
" minima....	24.974±0.139	2.407±0.098	9.65	24.686±0.129	2.227±0.091	9.02	24.368±0.120	2.068±0.085	8.49
Difference.......	+0.744±0.197	−0.005±0.139	−0.30	+1.080±0.182	−0.008±0.129	−0.41	+1.411±0.168	−0.038±0.119	−0.62
Diff./E_{diff}....	3.78	0.04		5.93	0.07		8.40	0.32	
Percentage difference........	2.94	0.21		4.28	0.38		5.63	1.88	
Calories per square meter.									
From means....	925.147±3.877	67.038±2.742	7.25	921.715±3.753	64.891±2.654	7.04	916.272±3.655	63.190±2.584	6.90
" minima....	898.926±3.983	68.870±2.817	7.66	883.777±3.955	68.382±2.797	7.74	866.893±3.879	67.067±2.743	7.74
Difference.......	+26.221±5.558	−1.832±3.931	−0.41	+37.938±5.452	−3.491±3.856	−0.70	+49.379±5.330	−3.877±3.768	−0.84
Diff./E_{diff}....	4.72	0.47		6.96	0.91		9.26	1.03	
Percentage dif-									

The fact that the average metabolism is lower when it is calculated from individual minima than when it is computed from individual means furnishes no argument in favor of either of the methods of computing the metabolism constant. Conclusions in regard to this point must be drawn from the results for weighting discussed above, and from a consideration of the variabilities.

From Table X we note that in whatever units heat production is expressed, the variation in the population metabolism (measured in either the absolute terms of the standard deviation or in the relative terms of the coefficient of variation) is lower when the individual means are employed than when individual minima are used as a basis for calculating the population constants.

If the securing of a constant with the lowest probable error is one of the goals to be attained, the method of means is, therefore, to be preferred over the method of minima.

III. SUMMARY.

In all special investigations in human calorimetry some standard constant measuring the metabolism of the normal individual must be used as a basis of comparison. The selection of this constant presents a problem of considerable difficulty from three sides.

The first is that of the physiological conditions under which the basal metabolism of the individual shall be measured; the second is that of the unit in which the caloric output of the individual shall be expressed; the third is that of the method by which the statistical constants for the standard series shall be obtained.

Basal metabolism measurements are generally made in two or more periods, with the subject in the postabsorptive state and in complete muscular repose, on the same day. Experimental periods which show evidence of muscular activity or of faulty technique in the analyses are discarded. So called minimum periods are utilized for obtaining a mean for the day. This may be designated as the daily mean.

It may be reasonably assumed that the results of the several periods of measurement on a given day stand in the relation of duplicate, triplicate, etc. analyses, and that it is not necessary to

ON THE OSMOTIC CONCENTRATION OF THE TISSUE FLUIDS OF PHANEROGAMIC EPIPHYTES

J. Arthur Harris

ON THE OSMOTIC CONCENTRATION OF THE TISSUE
FLUIDS OF PHANEROGAMIC EPIPHYTES

ON THE OSMOTIC CONCENTRATION OF THE TISSUE FLUIDS OF PHANEROGAMIC EPIPHYTES[1]

J. ARTHUR HARRIS

INTRODUCTORY REMARKS

The purpose of this paper, which is one of a series dealing with the problem of the physico-chemical properties of vegetable saps in relation to environmental factors and to geographical distribution, is to present the results of three series of determinations of the osmotic concentration of the tissue fluids of phanerogamic epiphytes, and to compare them briefly and in a preliminary way with available data for the osmotic concentrations found in the sap of terrestrial vegetation.

Notwithstanding the enthusiastic interest aroused in the mind of the botanical traveler by the remarkable range of form and the obvious physiological peculiarities of the Orchidaceae, Bromeliaceae, and other epiphytic forms so characteristic of tropical vegetation, our knowledge, in quantitative terms, of the physiology of these organisms is exceedingly meager.

Since I hope on another occasion to discuss epiphytism in greater detail, I shall not in this place review the general literature.

MATERIALS AND METHODS

In this paper I have meant to include only those species which may unquestionably be considered typical epiphytes. It was for this reason that a few determinations made on plants which may be either terrestrial or epiphytic were included by Mr. Lawrence and myself in our paper on the Jamaican montane rain forest vegetation (1917a). In some instances it is extremely difficult to determine just which species shall be regarded as epiphytes. Our data are given in detail, and any botanist who chooses may arrange them differently.

The methods employed in the present study are those sufficiently described in our earlier discussion of the parasitic and the terrestrial vegetation of the Blue Mountains (Harris and Lawrence, 1916, 1917a).

The determinations here recorded were secured in three periods of field work, the first in Jamaica in 1915, the second and third in southern Florida in 1916 and 1917.   In the first period I had the advantage of the co-operation of Mr. John V. Lawrence, who remained on the island for some time longer than I was able to do, and to whom I am indebted for a large part of the work on Jamaican forms.   In the third period Mr. Charles W. Crane rendered most efficient service in several phases of the work.

The determinations were carried out in the Tropical Laboratory at Cinchona, Jamaica, and in the Subtropical Laboratory of the United States Department of Agriculture at Miami, Florida.   I have to thank Mr. William Harris, F.L.S., and the members of the British Association Committee for the use of the Laboratory at Cinchona, and am much indebted to Dr. David Fairchild, Agricultural Explorer, and to Mr. Edward Simmonds, in charge of the Plant Introduction Garden at Miami, for the use of the laboratory and other favors.   All the species were determined in the herbarium of the New York Botanical Garden. In addition, I am indebted to Dr. Small for various courtesies in the field work.

PRESENTATION OF DATA

The following protocol gives the individual determinations for the several species in terms of freezing point lowering, Δ, corrected for undercooling, and osmotic concentration in atmospheres as determined from a published table (Harris and Gortner, 1914).   The averages, designated by bars for each species, are given at the extreme right. When only a single determination is available it has of necessity served to represent the species in place of the average.

In the Bromeliaceae an attempt has been made to arrange the forms in a rough series from the most typical tank forms to those departing most widely from the type in which water storage in the bases of the leaves is possible.   Ultimately I hope our determinations will cover a range of forms sufficiently wide and be numerous enough to justify consideration of the problem of the relationship between sap properties and morphological structure in this fascinating family of plants.   It has not seemed feasible to attempt any logical classification of the Orchidaceae, and they are merely alphabetically arranged for each of the regions.

All the Jamaican montane rain forest determinations were made in 1915.   Hence the year is omitted when dates are cited.   In the case of

the Florida determinations, the year as well as the day of the month has been given.

BROMELIACEAE

Guzmannia Sintensii (Baker) Mez $\overline{\Delta} = 0.31, \overline{P} = 3.8$

Montane Rain Forest, Leeward Slopes, Feb. 24, $\Delta = 0.31, P = 3.7$; Ridges, Feb. 9, $\Delta = 0.34$, $P = 4.1$; Mar. 9, $\Delta = 0.45$, $P = 5.5$; Jim Crow Peak, Feb. 17, $\Delta = 0.25$, $P = 3.0$; Feb. 17, $\Delta = 0.28$, $P = 3.4$; Windward Slopes and Ravines, Feb. 13, $\Delta = 0.25$, $P = 3.0$; Feb. 13, $\Delta = 0.23$, $P = 2.8$; Feb. 20, $\Delta = 0.28$, $P = 3.3$; Mar. 13, $\Delta = 0.44$, $P = 5.3$.

Guzmannia capituligera (Griseb.) Mez (?) $\overline{\Delta} = 0.44, \overline{P} = 5.2$

Montane Rain Forest, Leeward Slopes, Feb. 18[2], $\Delta = 0.46$, $P = 5.5$; Windward Slopes and Ravines, Feb. 22, $\Delta = 0.41$, $P = 4.9$.

Guzmannia monostachya (L.) Rusby $\overline{\Delta} = 0.46, \overline{P} = 5.6$

Subtropical Florida. Sykes Hammock, Jan. 27, 1916, $\Delta = 0.42$, $P = 5.1$; Mar. 16, 1917, $\Delta = 0.50$, $P = 6.0$.

Catopsis Berteroniana (Schult.) Mez $\overline{\Delta} = 0.46, \overline{P} = 5.6$

Subtropical Florida. Hattie Bauer Hammock, Mar. 19, 1917, $\Delta = 0.46$, $P = 5.5$; Mar. 19, 1917, $\Delta = 0.42$, $P = 5.0$; Royal Palm Hammock, Mar. 3, 1916, $\Delta = 0.50$, $P = 6.0$; Feb. 21, 1917, $\Delta = 0.48$, $P = 5.7$; Small Hammock between Florida City and Biscayne Bay, Feb. 17, 1917, $\Delta = 0.46$, $P = 5.6$.

Tillandsia utriculata L. $\overline{\Delta} = 0.43, \overline{P} = 5.2$

Subtropical Florida. Hattie Bauer Hammock, Mar. 16, 1917, $\Delta = 0.43$, $P = 5.2$; Mar. 19, 1917, $\Delta = 0.45$, $P = 5.4$; Royal Palm Hammock, Mar. 4, 1916, $\Delta = 0.42$, $P = 5.1$; Mar. 3, 1916, $\Delta = 0.33$, $P = 4.0$; Small Hammock near Royal Palm Hammock, Feb. 23, 1917, $\Delta = 0.43$, $P = 5.2$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.37$, $P = 4.5$; Palm and Live Oak Hammock, Peninsula, near the Narrows, Indian River, Apr. 1, 1917, $\Delta = 0.44$, $P = 5.3$; on dwarfed *Rhizophora Mangle*, near Biscayne Bay, Feb. 17, 1917, $\Delta = 0.40$, $P = 4.8$; Feb. 17, 1917, $\Delta = 0.57$, $P = 6.9$; Feb. 17, 1917, $\Delta = 0.39$, $P = 4.7$; Feb. 17, 1917, $\Delta = 0.34$, $P = 4.1$; Orange Grove, Miami, Feb. 10, 1917, $\Delta = 0.58$, $P = 6.9$.

[2] These determinations are for the older, outer leaves. Sap from the younger leaves gave $\Delta = 0.35$, $P = 4.2$.

Tillandsia Valenzuelana A. Rich. $\overline{\Delta} = 0.36, \overline{P} = 4.3$

Subtropical Florida. Royal Palm Hammock, Jan. 29, 1916, $\Delta = 0.35$, $P = 4.2$; Jan. 29, 1916, $\Delta = 0.36$, $P = 4.3$; Feb. 23, 1917, $\Delta = 0.34$, $P = 4.0$; Small Hammock near Royal Palm Hammock, Feb. 23, 1917, $\Delta = 0.41$, $P = 4.9$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.32$, $P = 3.9$; Brickell Hammock, Mar. 9, 1917, $\Delta = 0.38$, $P = 4.5$.

Tillandsia incurva Griseb. (?) $\overline{\Delta} = 0.25, \overline{P} = 3.0$

Montane Rain Forest, Ridges, Feb. 18, $\Delta = 0.28$, $P = 3.3$; Windward Slopes and Ravines, Mar. 2, $\Delta = 0.22$, $P = 2.7$.

Tillandsia fasciculata Swartz $\overline{\Delta} = 0.39, \overline{P} = 4.6$

Subtropical Florida. Hattie Bauer Hammock, Mar. 16, 1917, $\Delta = 0.36$, $P = 4.3$; Mar. 19, 1917, $\Delta = 0.33$, $P = 3.9$; Royal Palm Hammock, Mar. 4, 1916, $\Delta = 0.40$, $P = 4.8$; Feb. 27, 1917, $\Delta = 0.41$, $P = 5.0$; Small Pineland Hammock, near Royal Palm Hammock, Mar. 4, 1916, $\Delta = 0.44$, $P = 5.3$; Sykes Hammock, Jan. 27, 1916, $\Delta = 0.40$, $P = 4.8$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.31$, $P = 3.7$; Small Hammock near Royal Palm Hammock, Feb. 23, 1917, $\Delta = 0.43$, $P = 5.2$; Murden Hammock, Jan. 28, 1916, $\Delta = 0.45$, $P = 5.4$; Small Hammock between Biscayne Bay and Florida City, Feb. 17, 1917, $\Delta = 0.32$, $P = 3.9$.

Tillandsia aloifolia Hook.

Subtropical Florida. On dwarfed *Rhizophora Mangle* near Biscayne Bay, Feb. 17, 1917, $\Delta = 0.43$, $P = 5.1$.

Tillandsia Balbisiana Schult. $\overline{\Delta} = 0.46, \overline{P} = 5.5$

Subtropical Florida. Hattie Bauer Hammock, Mar. 19, 1917, $\Delta = 0.53$, $P = 6.4$; Royal Palm Hammock, Jan. 29, 1916, $\Delta = 0.40$, $P = 4.8$; Mar. 3, 1916, $\Delta = 0.39$, $P = 4.7$; Feb. 23, 1917, $\Delta = 0.47$, $P = 5.7$; Small Hammock near Royal Palm Hammock, Feb. 23, 1917, $\Delta = 0.49$, $P = 6.0$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.38$, $P = 4.6$; on dwarfed *Rhizophora Mangle* near Biscayne Bay, Feb. 29, 1916, $\Delta = 0.50$, $P = 6.0$; Feb. 17, 1917, $\Delta = 0.51$, $P = 6.1$.

Tillandsia tenuifolia L. $\overline{\Delta} = 0.42, \overline{P} = 5.1$

Subtropical Florida. Royal Palm Hammock, Jan. 29, 1916, $\Delta = 0.38$, $P = 4.6$; Feb. 21, 1917, $\Delta = 0.45$, $P = 5.4$; Sykes Hammock, Jan. 27, 1916, $\Delta = 0.46$, $P = 5.5$; Mar. 15, 1917, $\Delta = 0.45$, $P = 5.5$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.37$, $P = 4.4$.

Tillandsia recurvata L. $\overline{\Delta} = 0.49, \overline{P} = 5.8$

Subtropical Florida. Royal Palm Hammock, Mar. 3, 1916, $\Delta = 0.52$, $P = 6.2$; Feb. 23, 1917, $\Delta = 0.45$, $P = 5.4$.

Dendropogon usneoides (L.) Raf. $\overline{\Delta} = 0.75, \overline{P} = 9.0$

Subtropical Florida. Hattie Bauer Hammock, Mar. 19, 1917, $\Delta = 0.62$, $P = 7.5$; Royal Palm Hammock, Mar. 3, 1916, $\Delta = 0.86$, $P = 10.4$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.50$, $P = 6.0$; Palm and Live Oak Hammock, the Peninsula near the Narrows, Indian River, Apr. 1, 1917, $\Delta = 1.32$, $P = 15.8$; Sykes Hammock, Mar. 15, 1917, $\Delta = 0.57$, $P = 6.8$; Orange Grove, Miami, Mar. 6, 1916, $\Delta = 0.70$, $P = 8.4$; Feb. 9, 1917, $\Delta = 0.65$, $P = 7.9$.

ORCHIDACEAE

Epidendrum imbricatum Lindl.

Montane Rain Forest, Windward Slopes and Ravines, Feb. 13, $\Delta = 0.30$, $P = 3.6$.

Lepanthes ovalis (Swartz) Fawc. & Rendle

Montane Rain Forest, Leeward Ravines, Mar. 18, $\Delta = 0.35$, $P = 4.2$.

Lepanthes divaricata Fawc. & Rendle $\overline{\Delta} = 0.20, \overline{P} = 2.4$

Montane Rain Forest, Ridges, Feb. 9, $\Delta = 0.19$, $P = 2.3$; Mar. 9, $\Delta = 0.27$, $P = 3.3$; Jim Crow Peak, Feb. 17, $\Delta = 0.16$, $P = 1.9$; Windward Slopes and Ravines, Feb. 20, $\Delta = 0.19$, $P = 2.3$; Feb. 24, $\Delta = 0.18$, $P = 2.2$; Mar. 4, $\Delta = 0.20$, $P = 2.4$.

Octadesmia montana (Swartz) Benth. $\overline{\Delta} = 0.44, \overline{P} = 5.3$

Montane Rain Forest, Jim Crow Peak, Feb. 17, $\Delta = 0.41$, $P = 5.0$; Windward Slopes and Ravines, Feb. 20, $\Delta = 0.46$, $P = 5.5$.

Pleurothallis racemiflora (Swartz) Lindl.

Montane Rain Forest, Leeward Ravines, Mar. 11, $\Delta = 0.21$, $P = 2.6$.

Stelis micrantha Swartz $\overline{\Delta} = 0.22, \overline{P} = 2.6$

Montane Rain Forest, Ridges, Feb. 9, $\Delta = 0.23$, $P = 2.7$; Windward Slopes and Ravines, Feb. 4, $\Delta = 0.24$, $P = 2.9$; Feb. 13, $\Delta = 0.20$, $P = 2.4$; Feb. 20, $\Delta = 0.21$, $P = 2.5$.

Stelis ophioglossoides Swartz

Montane Rain Forest, Jim Crow Peak, Feb. 17, $\Delta = 0.22$, $P = 2.7$.

Anacheilium cochleatum (L.) Hoffmannsegg $\overline{\Delta}$ = 0.43, $\overline{P}$ = 5.2

Subtropical Florida. Hattie Bauer Hammock, Jan. 28, 1916, Δ = 0.44, P = 5.3; Royal Palm Hammock, Jan. 29, 1916, Δ = 0.41, P = 5.0.

Auliza nocturna (L.) Small $\overline{\Delta}$ = 0.42, $\overline{P}$ = 5.0

Subtropical Florida. Hattie Bauer Hammock, Jan. 28, 1916, Δ = 0.46, P = 5.5; Mar. 16, 1917, Δ = 0.47, P = 5.7; Mar. 19, 1917, Δ = 0.52, P = 6.3; Royal Palm Hammock, Jan. 29, 1916, Δ = 0.38, P = 4.5; Feb. 21, 1917, Δ = 0.29, P = 3.5; Small Pineland Hammock, near Royal Palm Hammock, Mar. 3, 1916, Δ = 0.41, P = 4.9; Bryan Hammock, Feb. 13, 1917, Δ = 0.40, P = 4.8; Feb. 13, 1917, Δ = 0.39, P = 4.6.

Encyclia tampense (Lindl.) Small $\overline{\Delta}$ = 0.48, $\overline{P}$ = 5.8

Subtropical Florida. Hattie Bauer Hammock, Jan. 28, 1916, Δ = 0.48, P = 5.8; Mar. 16, 1917, Δ = 0.50, P = 6.1; Mar. 19, 1917, Δ = 0.51, P = 6.2; Royal Palm Hammock, Jan. 29, 1916, Δ = 0.44, P = 5.3; Brickell Hammock, Mar. 22, 1917, Δ = 0.49, P = 5.9; Mar. 24, 1917, Δ = 0.40, P = 4.8; Bryan Hammock, Feb. 13, 1917, Δ = 0.45, P = 5.4; Palm and Live Oak Hammock, Peninsula, near the Narrows, Indian River, Apr. 1, 1917, Δ = 0.62, P = 7.4; Small Hammock near Royal Palm Hammock, Feb. 23, 1917, Δ = 0.47, P = 5.6.

Macradenia lutescens R. Br. $\overline{\Delta}$ = 0.51, $\overline{P}$ = 6.1

Subtropical Florida. Royal Palm Hammock, Jan. 29, 1916, Δ = 0.53, P = 6.4; Feb. 21, 1917, Δ = 0.48, P = 5.7.

Polystachya minuta (Aubl.) Britton

Subtropical Florida. Bryan Hammock, Feb. 13, 1917, Δ = 0.50, P = 6.0.[3]

Spathiger rigidus (Jacq.) Small $\overline{\Delta}$ = 0.38, $\overline{P}$ = 4.5

Subtropical Florida. Hattie Bauer Hammock, Jan. 28, 1916, Δ = 0.47, P = 5.7; Mar. 16, 1917, Δ = 0.36, P = 4.4; Mar. 19, 1917, Δ = 0.42, P = 5.0; Royal Palm Hammock, Jan. 29, 1916, Δ = 0.35,

[3] Sample from Hattie Bauer Hammock obtained January 28, 1916, and one from the Brickell Hammock, March 22, 1917, were so mucilaginous that no determination could be made. The juice of the sample from the Bryan Hammock was also highly mucilaginous and could not be filtered. Until verification this determination must be taken as only approximate.

$P = 4.2$; Jan. 29, 1916, $\Delta = 0.31$, $P = 3.7$; Mar. 3, 1916, $\Delta = 0.40$, $P = 4.8$; Feb. 21, 1917, $\Delta = 0.32$, $P = 3.9$.

Vanilla Eggersii Rolfe

Subtropical Florida. Brickell Hammock, Feb. 14, 1916, $\Delta = 0.24$, $P = 2.9$. This determination is of course based on sap from the stems.

PIPERACEAE

Peperomia basellifolia H.B.K. $\qquad \overline{\Delta} = 0.35,\ \overline{P} = 4.2$

Montane Rain Forest, Windward Slopes and Ravines, Feb. 20, $\Delta = 0.40$, $P = 4.8$; Feb. 24, $\Delta = 0.35$, $P = 4.2$; Mar. 4, $\Delta = 0.33$, $P = 3.9$; Mar. 13, $\Delta = 0.31$, $P = 3.7$.

Peperomia crassicaulis Fawc. & Rendle $\qquad \overline{\Delta} = 0.40,\ \overline{P} = 4.9$

Montane Rain Forest, Ridges, Feb. 18, $\Delta = 0.40$, $P = 4.8$; Mar. 9, $\Delta = 0.44$, $P = 5.2$; Mar. 13, $\Delta = 0.42$, $P = 5.1$; Mar. 16, $\Delta = 0.46$, $P = 5.6$; Windward Slopes and Ravines, Mar. 4, $\Delta = 0.30$, $P = 3.6$.

Peperomia magnoliafolia (Jacq.) A. Dietr. $\qquad \overline{\Delta} = 0.38,\ \overline{P} = 4.6$

Subtropical Florida, Royal Palm Hammock, Jan. 29, 1916, $\Delta = 0.38$, $P = 4.6$; Small Hammock near Royal Palm Hammock, Feb. 23, 1917, $\Delta = 0.39$, $P = 4.7$; Bryan Hammock, Feb. 13, 1917, $\Delta = 0.37$, $P = 4.5$; Feb. 13, 1917, $\Delta = 0.35$, $P = 4.2$; Sykes Hammock, Jan. 27, 1916, $\Delta = 0.41$, $P = 4.9$.

Peperomia Myrtillus Miquel $\qquad \overline{\Delta} = 0.36,\ \overline{P} = 4.3$

Montane Rain Forest, Leeward Ravines, Mar. 11, $\Delta = 0.35$, $P = 4.2$; Windward Slopes and Ravines, Mar. 13, $\Delta = 0.36$, $P = 4.4$.

Peperomia quadrifolia (L.) H.B.K.

Montane Rain Forest, Leeward Ravines, Mar. 11, $\Delta = 0.39$, $P = 4.6$.

Peperomia septemnervis Ruiz & Pav. $\qquad \overline{\Delta} = 0.31,\ \overline{P} = 3.7$

Montane Rain Forest, Leeward Ravines, Mar. 11, $\Delta = 0.33$, $P = 3.9$; Windward Slopes and Ravines, Feb. 13, $\Delta = 0.31$, $P = 3.7$; Feb. 13, $\Delta = 0.30$, $P = 3.6$.

GESNERACEAE

Columnea hirsuta Swartz $\qquad \overline{\Delta} = 0.36,\ \overline{P} = 4.3$

Montane Rain Forest, Leeward Ravines, Feb. 26, $\Delta = 0.40$, $P = 4.8$; Windward Slopes and Ravines, Feb. 13, $\Delta = 0.33$, $P = 4.0$; Feb. 20, $\Delta = 0.34$, $P = 4.1$; Feb. 22, $\Delta = 0.33$, $P = 4.0$; Feb .22,

$\Delta = 0.32$, $P = 3.9$; Feb. 24, $\Delta = 0.36$, $P = 4.4$; Mar. 2, $\Delta = 0.37$, $P = 4.5$; Mar. 4, $\Delta = 0.38$, $P = 4.5$; Mar. 13, $\Delta = 0.40$, $P = 4.8$.

ANALYSIS OF DATA

In this paper I shall limit discussion of the data presented to a comparison of the constants of the epiphytes among themselves and with the values which have already been obtained for terrestrial forms in various habitats. Even these comparisons must be limited by the still unorganized condition of our data for several important habitats. Since, however, it will be many months before all of these data can be fully analyzed and ready for discussion, it has seemed proper to place the data which have been obtained for epiphytes during the past three years on record in a form which will enable other physiologists and phytogeographers to use them.

Consider first of all the relative magnitudes of the osmotic concentrations found in the epiphytic plants of the two regions considered. The results, grouped by families, are shown in table 1.

The constants in this table are the averages of species means, not of species determinations (except when only one determination is available for a species), for each family. While the species means which are based upon a large number of determinations are somewhat more trustworthy than those which are based upon only two or three readings, or upon only a single collection, the general mean for the habitat is certainly more representative when calculated in this way than if the habitat average had been computed directly from the individual constants, thus weighting the species with the numbers of collections of each which happened to be made.

TABLE I

Comparison of Osmotic Concentrations in Jamaican and Floridian Epiphytes

	Jamaica	Florida	Difference
Bromeliaceae.....	$\overline{\Delta} = 0.333$, $\overline{P} = 4.00$ 2 genera, 3 species	$\overline{\Delta} = 0.464$, $\overline{P} = 5.57$ 4 genera, 10 species	$\overline{\Delta} = +0.131$, $\overline{P} = +1.57$
Orchidaceae......	$\overline{\Delta} = 0.276$, $\overline{P} = 3.32$ 5 genera, 7 species	$\overline{\Delta} = 0.421$, $\overline{P} = 5.06$ 7 genera, 7 species	$\overline{\Delta} = +0.145$, $\overline{P} = +1.74$
Piperaceae.......	$\overline{\Delta} = 0.362$, $\overline{P} = 4.34$ Peperomia only, 5 species	$\overline{\Delta} = 0.380$, $\overline{P} = 4.58$ *Peperomia magnolia-folia* only	$\overline{\Delta} = +0.018$, $\overline{P} = +0.24$
Gesneraceae......	$\overline{\Delta} = 0.358$, $\overline{P} = 4.33$ *Columnea hirsuta* only	—————— No representative.	——————

The table brings out clearly two facts:

1. That in all four families and in both Jamaica and Florida, the osmotic concentration of epiphytic forms is extremely low.

2. That for the three groups represented in both regions the osmotic concentration of the epiphytes (chiefly from the hammocks) of subtropical Florida is higher than that demonstrated in the Jamaican rain forest. The average difference is 1.57 atmospheres higher for the Bromeliaceae,[4] 1.74 atmospheres higher for the Orchidaceae, and 0. 24 atmospheres higher for the single species of *Peperomia*.

The comparison may be made somewhat more analytically on the basis of the means for the genera.

The constants in table 2 are averages of the species means of each of the genera.

TABLE 2

Genera of Jamaican and Floridian Epiphytes Arranged in the Order of the Average Osmotic Concentration of Their Species

Jamaica		Florida	
Genus	$\bar{P}$	$\bar{P}$	Genus
Pleurothallis.............	2.57		
Stelis...................	2.65		
		2.90	Vanilla
Tillandsia...............	3.00		
Lepanthes...............	3.28		
Epidendrum..............	3.56		
Columnea...............	4.33		
Peperomia...............	4.34		
Guzmannia..............	4.49		
		4.52	Spathiger
		4.58	Peperomia
		4.97	Auliza
		5.09	Tillandsia
		5.15	Anacheilium
Octadesmia..............	5.25		
		5.55	Guzmannia
		• 5.56	Catopsis
		5.83	Encyclia
		6.00	Polystachya
		6.05	Macradenia
		8.97	Dendropogon

[4] That the higher value for Floridian Bromeliaceae is not primarily due to the inclusion of *Dendropogon usneoides* (= *Tillandsia usneoides*) is shown by the fact that if this species be omitted from the Florida series, the remaining 9 species average $\bar{\Delta} = 0.433$, $\bar{P} = 5.19$, which are respectively 0.100 and 1.19 greater than the Jamaican average.

It is clear at a glance that with the exception of the stem-succulent Vanilla in the Floridian and of *Octadesmia montana* in the Jamaican constants, the two series do not overlap in the average (generic) magnitude of their constants. With the exceptions noted, the Jamaican (rain forest) genera range from 2.57 to 4.49 atmospheres, whereas the Floridian genera range from 4.52 to 8. 97 atmospheres.

Instead of limiting our comparisons between the two regions to means, the individual determinations may be seriated according to their magnitude and the frequency distributions compared. This has the advantage of giving a general view of the range of variation in the individual constants, but the disadvantage from the standpoint of exact comparison that certain species are far more extensively represented than others. The frequency distributions are given in table 3.

TABLE 3

Frequency Distributions of Osmotic Concentration Determinations in Jamaican and Floridian Epiphytes

Osmotic Concentration in Atmospheres	Orchidaceae		Bromeliaceae	
	Jamaica	Florida	Jamaica	Florida
1.5–1.9	1	—	—	—
2.0–2.4	5	—	—	—
2.5–2.9	5	1	2	—
3.0–3.4	1	—	5	—
3.5–3.9	1	3	1	4
4.0–4.4	1	2	1	7
4.5–4.9	—	6	1	11
5.0–5.4	1	5	1	14
5.5–5.9	1	7	2	6
6.0–6.4	—	5	—	7
6.5–6.9	—	—	—	2
7.0–7.4	—	1	—	—
	16	30	13	51

Because of the unusually high values found in the Spanish moss (*Dendropogon usneoides*) it has been omitted from this table. Notwithstanding this fact, the Floridian Bromeliaceae as well as the Orchidaceae show distinctly higher minima and maxima than the Jamaican forms. The distinction between the two regions is not as clearly shown by the distribution of the individual determinations as by the generic means, since individual determinations must be expected to show much wider variation than averages.

I now turn to the relative magnitude of the osmotic concentration of terrestrial and epiphytic plants.

Since in a number of series of determinations we have found a differentiation in the sap properties of ligneous and herbaceous plants,[5] I shall compare epiphytic Orchidaceae, Bromeliaceae, and Piperaceae primarily with terrestrial herbaceous plants.

Unfortunately the several hundreds of determinations from the various coastal, pineland, hammock, and Everglade habitats of Subtropical Florida are as yet unclassified, and it will probably require some time before the results from this highly interesting region are discussed in detail.

The averages for the various groups of epiphytes from Jamaica and from Subtropical Florida have been given in table 1.

The average freezing-point lowering of the saps ranges from 0.276° to 0.464°, less than two tenths of one degree. In terms of osmotic concentration the values lie between 3.3 and 5.6 atmospheres, a range of less than two and one third atmospheres.

The only extensive series of averages for herbaceous terrestrial vegetation are those for the Arizona deserts made by Harris, Lawrence, and Gortner (1916), and the first Long Island series, by Harris, Lawrence, and Gortner, as yet unpublished, and the Jamaican montane rain forest series which will be treated in greater detail below.

For the Long Island habitats the preliminary average values are:

Habitat	Average Concentration, $\bar{P}$
Beaches, coastal sand dunes, and marshes	13.62
Dryer woods and open fields	10.04
Permanently moist localities	9.27
All habitats	10.41

Note that the epiphytic forms show a sap concentration about one third to one half as great.

For the Arizona desert (vernal) flora the averages for herbaceous plants are:

[5] For averages for divers growth forms from the Arizona deserts see Harris, Lawrence, and Gortner (1916). Averages for Long Island and Jamaican habitats are given by Harris and Lawrence (1917a). Some general comparisons are made by Harris (1917).

Habitat	Average Concentration, $\bar{P}$
Rocky slopes	15.94
Canyons	13.33
Arroyos	12.99
Bajada slopes	20.53
Salt spots	23.57
All habitats	15.15

These values are (roughly speaking) from 4 to 7 times as large as those for the epiphytic families.

For the Jamaican series, and unfortunately only for the Jamaican series, it is possible at this time to compare the averages for epiphytic and terrestrial forms from the same habitat.

Table 4 gives the averages of the species means for each habitat for the Orchidaceae and Bromeliaceae and for the genus Peperomia of

TABLE 4

Comparison of Osmotic Concentration of Epiphytic Plants with that of Terrestrial Herbaceous Plants in the Montane Rain Forest

Habitats	Average for Terrestrial Herbs	Orchidaceae — Average for Epiphytic Orchidaceae	Orchidaceae — Difference and Relative Value	Bromeliaceae — Average for Epiphytic Bromeliaceae	Bromeliaceae — Difference and Relative Value	Piperaceae — Average for Epiphytic Piperaceae	Piperaceae — Difference and Relative Value
Ruinate of the leeward slopes	$\bar{\Delta}=0.812$ $\bar{P}=9.77$ $(n=17)$	—	—	$\bar{\Delta}=0.385$ $\bar{P}=4.60$ $(n=2)$	-0.427 -5.17 47.1%	—	—
Leeward ravines	$\bar{\Delta}=0.628$ $\bar{P}=7.59$ $(n=13)$	$\bar{\Delta}=0.280$ $\bar{P}=3.40$ $(n=2)$	-0.348 -4.19 44.8%	—	—	$\bar{\Delta}=0.357$ $\bar{P}=4.23$ $(n=3)$	-0.271 -3.36 55.7%
Ridges and peaks	$\bar{\Delta}=0.718$ $\bar{P}=8.63$ $(n=8)$	$\bar{\Delta}=0.267$ $\bar{P}=3.22$ $(n=4)$	-0.451 -5.41 37.3%	$\bar{\Delta}=0.305$ $\bar{P}=3.65$ $(n=2)$	-0.413 -4.98 42.3%	$\bar{\Delta}=0.431$ $\bar{P}=5.19$ $(n=1)$	-0.287 -3.44 60.1%
Windward slopes and ravines	$\bar{\Delta}=0.627$ $\bar{P}=7.52$ $(n=15)$	$\bar{\Delta}=0.292$ $\bar{P}=3.50$ $(n=4)$	-0.335 -4.02 46.5%	$\bar{\Delta}=0.310$ $\bar{P}=3.73$ $(n=3)$	-0.317 -3.79 49.6%	$\bar{\Delta}=0.330$ $\bar{P}=3.98$ $(n=4)$	-0.297 -3.54 52.9%
All habitats	$\bar{\Delta}=0.700$ $\bar{P}=8.43$ $(n=53)$	$\bar{\Delta}=0.280$ $\bar{P}=3.37$ $(n=10)$	-0.420 -5.06 40.0%	$\bar{\Delta}=0.330$ $\bar{P}=3.96$ $(n=7)$	-0.370 -4.47 47.0%	$\bar{\Delta}=0.353$ $\bar{P}=4.23$ $(n=8)$	-0.347 -4.20 50.2%

the Piperaceae. The number under each of the averages is the number of species, not the number of determinations, upon which it is based.

The averages for terrestrial herbaceous species are those already published (Harris and Lawrence, 1917a).

The general mean for the region has been computed by averaging the species means for the individual habitats. Thus if a species occurs in both the Leeward Ravines and the Ridge Forest it is counted twice, whereas the species which occur in one of these habitats only will be counted but once. Thus the numbers of the species given for all habitats is the number of species weighted with the number of the sub-habitats in which they occur.[6]

The comparison between the epiphytic and the terrestrial herbaceous forms has been made in two ways. First, the actual differences in the average depression of the freezing point and in the average calculated osmotic concentration have been determined and are given with their signs. Second, the average values of P of the epiphytes have been expressed as a percentage of the value for terrestrial herbs.[7]

An examination of the nine comparisons between the epiphytic and terrestrial herbs of the four individual habitats shows that the concentration is in every instance lower for the epiphytic forms. The averages are roughly 4.0 to 5.4 atmospheres lower in the Orchidaceae, 3.8 to 5.2 atmospheres lower in the Bromeliaceae, and 3.4 to 3.6 atmospheres lower in Peperomia of the Piperaceae.[8]

There now remains for consideration only the half shrubby gesneraceous epiphyte *Columnea hirsuta*. One determination from the Leeward Ravines gives $\Delta = 0.395$, $P = 4.76$. Eight constants from the Windward Slopes and Ravine average $\overline{\Delta} = 0.354$, $\overline{P} = 4.28$. If these be compared with the averages for herbaceous vegetation from the same habitats, differences in P of -2.83 for the Leeward Ravine determination and of -3.24 for the Windward habitats are secured.

[6] This method of computing the average has both advantages and disadvantages. For present purposes it is quite adequate.

[7] Practically the same percentages are secured by using the average values of freezing-point lowering, but since the relationship between Δ and P is not strictly linear the results are not exactly identical.

[8] Comparisons with the herbaceous plants of the regions as a whole show a concentration 5.1 atmospheres lower for Orchidaceae, 4.5 atmospheres lower for Bromeliaceae, and 4.2 atmospheres lower for Peperomia of the Piperaceae. The averages for the whole region is obtained by weighting those of the individual habitats with the number of species examined.

If the comparison be made with the ligneous terrestrial vegetation, the differences are −6.07 for the Leeward and −5.45 for the Windward habitats.

In relative terms, the osmotic concentrations of the sap of the epiphytic Orchidaceae is only 37.3 to 46.5 percent as high as that of the terrestrial herbs of the same habitat, the constants for the Bromeliaceae range from 42.3 to 49.6 percent of the comparable values for terrestrial herbs, while the determinations based on Peperomia range from 52.9 to 60.1 percent of those for the non-epiphytic herbs of the same habitats. Columnea shows a concentration of 56.9 percent of that of herbaceous plants in the Windward habitats and 62.7 percent of that of herbaceous plants in the Leeward habitats. If compared with ligneous terrestrial vegetation it shows a concentration of 44.0 percent in the Windward and of 44.0 percent in the Leeward habits.

Summarizing the results of this comparison: the osmotic concentration of the fluids of the epiphytic Orchidaceae, Bromeliaceae, Piperaceae, and Gesneraceae of the montane rain forest of Jamaica is roughly speaking only 37.3 to 62.7 percent as high as that of the terrestrial plants of the same region.

The averages for herbaceous forms include, as already explained, a few determinations based on species which may occur on the ground or as epiphytes. They also include those based on a few ferns and fern allies. The removal of these constants might change *slightly* the actual values of the difference in the table. Since the forms which have been classified as terrestrial but may occur as epiphytes are characterized by lower osmotic concentration than the vegetation as a whole, the removal of these species from the list of herbaceous plants would make the differences demonstrated between terrestrial and epiphytic vegetation even larger. The exclusion of the few determinations for terrestrial ferns and fern allies could be justified only on the assumption that they are sensibly differentiated in their sap properties from flowering plants. There is, at present, no basis for such an assumption.

The low concentration of the sap of epiphytic Phanerogams may perhaps be most clearly brought out by comparing it with that of the ligneous species upon which they may occur. Table 5 gives the differences and relative concentrations for the Jamaican materials.

Epiphytic Orchidaceae show from 28 to 36 percent, the epiphytic Bromeliaceae from 32 to 38 percent, the epiphytic Piperaceae from 39 to 45 percent, and the epiphytic Gesneraceae about 44 percent

of the osmotic concentration exhibited by the foliage of the ligneous forms upon which they find lodgment. Of course these figures are only approximations, which will be somewhat modified by further work, but they are based on sufficient data to justify the conclusion that the epiphytic species of the rain forest are characterized by a concentration of about one third to one half that of the ligneous terrestrial species.

TABLE 5

Comparison of Osmotic Concentration of Epiphytic Forms with that of Ligneous Terrestrial Species in the Montane Rain Forest

Habitats	Average for Ligneous Plants	Difference and Relative Value			
		Orchidaceae	Bromeli- aceae	Piperaceae	Gesneraceae
Ruinate of the leeward slopes...	13.05 ($n=40$)	—	−8.45 35.2%	—	—
Leeward ravines..............	10.83 ($n=32$)	−7.43 31.4%	—	−6.60 39.1%	−6.07 44.0%
Ridges and peaks..:..........	11.54 ($n=36$)	−8.32 27.9%	−7.89 31.6%	−6.35 45.0%	—
Windward slopes and ravines...	9.73 ($n=28$)	−6.23 36.0%	−6.00 38.3%	−5.75 40.9%	−5.45 44.0%
All habitats.................	11.44 ($n=136$)	−8.07 29.5%	−7.48 34.6%	−7.21 37.0%	−7.11 37.8%

In passing it may be worth while to point out that these results have an important bearing upon theories of the origin of parasitism. The suggestion has been made that epiphytism is the first stage in the evolution of parasitism in the flowering plants. But all of these most typical epiphytes are characterized by very low osmotic concentration in comparison with the ligneous species of the same region, whereas the Loranthaceae of these forests have been shown (Harris and Lawrence, 1916) to have generally higher concentration of their tissue fluids than their hosts. Similar relationships have been found to exist in desert Loranthaceae (Harris, 1918).

Theoretically one of the best methods of comparison would be to lay side by side constants for terrestrial and epiphytic members of the same family. Unfortunately I have not been able to secure terrestrial Orchidaceae from subtropical Florida. Determinations have been published (Harris and Lawrence, 1917a) for Jamaican species. *Epidendrum verrucosum*, which we included in our first paper because we always found it growing on the ground, although Fawcett and Rendle

record it as occurring "on trees, rocks, and dry banks," gives on the average $\overline{\Delta} = 0.51$, $\overline{P} = 6.1$ in the Leeward ravines and $\overline{\Delta} = 0.55$, $\overline{P} = 6.6$ in the ruinate. *Prescottia stachyoides* from the windward ravines and slopes gave an average depression of $\overline{\Delta} = 0.52$, or an average concentration in atmospheres of $\overline{P} = 6.3$.

All of these values are distinctly, and in many cases very much, greater than those obtained from the individual species of epiphytic Orchidaceae.

For comparison with the epiphytic Peperomia we have only *Peperomia stellata*, which we collected in Jamaica only as a terrestrial herb. It gave the following values:

> Leeward ravines, $\overline{\Delta} = 0.43$, $\overline{P} = 5.2$
> Ridge Forest, $\overline{\Delta} = 0.45$, $\overline{P} = 5.4$
> Leeward habitats, $\overline{\Delta} = 0.42$, $\overline{P} = 5.1$

These values are slightly higher than the averages for any of the epiphytic species from the rain forest.

As far as I am aware, the only determination of osmotic concentration of the tissue fluids of any bromeliad hitherto made is that for *Bromelia Pinguim*, which Mr. Lawrence and I (1917*b*) found growing as a terrestrial plant in the Jamaican coastal deserts. This gave $\Delta = 0.63$, $P = 7.6$. This is a value higher than any of those recorded in this paper with the exception of those for *Dendropogon usneoides*. It is, however, extremely low for such a habitat as the Jamaican Coastal Deserts.

With regard to two species which Mr. Lawrence and I treated with the terrestrial vegetation but which others have observed growing as air plants, the following points may be noted.

The woody-stemmed *Blakea trinervia*, which may be rooted in the soil or, according to Shreve, grown as an epiphyte, has a concentration measured by $\overline{\Delta} = 0.58$, $\overline{P} = 6.9$, as compared with the general average of $\overline{\Delta} = 0.81$, $\overline{P} = 9.7$ for the ligneous species of the windward habitats in which it occurs.

Tradescantia multiflora, which we included with terrestrial vegetation in our earlier paper, but which may also occur as an epiphyte, gave in a single determination $\Delta = 0.39$, $P = 4.7$. This is far lower than the general averages of $\overline{\Delta} = 6.3$, $\overline{P} = 7.6$ for the herbs of the Leeward ravines.

CONCLUSIONS

The osmotic concentration of the tissue fluids of epiphytic Bromeliaceae, Orchidaceae, Piperaceae, and Gesneraceae is far lower than that of terrestrial vegetation.

In the Jamaican montane rain forest where direct comparisons for individual habitats are possible, the epiphytes show from 37 to 60 percent of the concentration characteristic of herbaceous terrestrial vegetation, and from 28 to 45 percent of the concentration of ligneous terrestrial vegetation.

In the Bromeliaceae, Orchidaceae, and Peperomia of the Piperaceae, the osmotic concentration of the species of the Jamaican montane rain forest is lower than that of the species of the hammocks of subtropical Florida.

At some future time I hope to deal with the problem of the osmotic concentration of cryptogamic epiphytes and to obtain data on the inorganic and organic constituents of the fluids of epiphytes which will justify further discussion of the physiology of these ecologically remarkable forms.

LITERATURE CITED

Harris, J. A. 1917. Physical chemistry in the service of phytogeography. Science, n. ser. **46**: 25–30.
——. 1918. On the osmotic concentration of the tissue fluids of desert Loranthaceae. Mem. Torrey Club **17**: 307–315.
Harris, J. A., and Gortner, R. A. 1914. Note on the calculation of the osmotic pressure of expressed vegetable saps from the depression of the freezing point, with a table for the values of P for $\Delta = .001°$ to $\Delta = 2.999°$. Amer. Journ. Bot. **1**: 75–78.
Harris, J. A., and Lawrence, J. V. 1916. On the osmotic pressure of the tissue fluids of Jamaican Loranthaceae parasitic on various hosts. Amer. Journ. Bot. **3**: 438–455. Diag. 1–2.
——. 1917a. The osmotic concentration of the tissue fluids of Jamaican montane rain forest vegetation. Amer. Journ. Bot. **4**: 268–298.
——. 1917b. Cryoscopic determinations on tissue fluids of plants of Jamaican coastal deserts. Bot. Gaz. **64**: 285–305.
——. With the co-operation of R. A. Gortner. 1916. The cryoscopic constants of expressed vegetable saps as related to local conditions in the Arizona deserts. Physiol. Res. **2**: 1–49.

ON THE RELATIONSHIP BETWEEN FREEZING POINT LOWERING, Δ, AND SPECIFIC ELECTRICAL CONDUCTIVITY, K, OF PLANT TISSUE FLUIDS

THE problem of the contribution of non-electrolytes, of undissociated molecules of electrolytes, and of dissociated ions of electrolytes to the depression of the freezing point, Δ, in terms of which osmotic concentration is usually measured, is one of considerable biological importance. We desire to know, for example, whether an observed difference in the osmotic concentration of the tissue fluids of a species growing in two different habitats is due primarily to differences in the quantities of electrolytes absorbed from the medium or to differences in the quantities of organic substances elaborated. The same question naturally arises when one is comparing the osmotic concentration of the tissue fluids of different species in the same habitat.

In the mixed solutions with which the biologist has to deal the problem presents serious difficulties. In certain cases some progress may be made by determining the correlation between the freezing point depression, Δ, and the specific electrical conductivity, K.

As a specific illustration we may take the relationship between osmotic concentration and electrical conductivity in a series of plant species growing in the non-halophytic habitats of the north shore of Long Island.[1]

In a series of 19 species of trees, 36 species of shrubs, and 162 species of herbs both Δ and

[1] Protocols of data and full details are given in a paper in press in the *Journal of Physical Chemistry*.

K are highly variable. The coefficients of variation, *i.e.*, 100 σ/m, where σ is the standard deviation and m the means are:

	Δ	K
Trees	21.46	28.49
Shrubs	18.46	28.03
Trees and shrubs ...	20.20	28.27
Herbs	23.46	25.33

Our problem is to determine whether higher values of K are associated with higher values of Δ, or whether within each of these growth forms[2] these two constants of the solution are essentially independent.

Determining the correlation coefficients by the usual product moment method we have the following measures of relationship between the magnitudes of K and Δ in the various series.

For trees, $N = 19$, $r = +0.127 \pm .152$
For shrubs, $N = 36$, $r = -0.079 \pm .112$
For trees and
shrubs, $N = 55$, $r = +0.022 \pm .091$
For herbs, $N = 162$, $r = +0.150 \pm .052$

For ligneous plants the correlations between Δ and K are low and statistically insignificant in comparison with their probable errors. The coefficient for shrubs is actually negative in sign. That for trees and shrubs together is sensibly zero. The coefficient for herbaceous plants is also low but may indicate a slight relationship between the two constants, higher values of Δ being associated with higher values of K and *vice versa*.

[2] It is necessary to separate the growth forms, since, as shown in detail elsewhere (Harris, Gortner and Lawrence, *loc. cit.*), the growth forms are highly differentiated with respect to both Δ and K. The actual means are:

	Δ	$K \times 10^6$
Trees	1.292	11,213
Shrubs	1.177	10,770
Trees and shrubs ..	1.217	10,923
Herbs	0.846	14,308

These results show that, in the vegetation of the glacial moraines of Long Island at least, there is practically no relationship between the concentration of ionized electrolytes and of total solutes (molecules and ions) in the leaf tissue fluids.[3]

J. Arthur Harris,
Ross Aiken Gortner,
John V. Lawrence

44 (1626)

The specific electrical conductivity of the tissue fluids of desert Loranthaceae.

By J. Arthur Harris and A. T. Valentine.

[*From the Station for Experimental Evolution, Cold Spring Harbor, L. I.*]

MacDougal and Cannon[1] and MacDougal[2] suggested some years ago that the osmotic concentration of the tissue fluids of the two organisms is one of the fundamental variables in the relationship between plant parasite and host. Senn[3] has published one plasmolytic determination indicating higher concentration in a *Viscum* than in the leaves of the host tree and has secured similar results with other phanerogamic parasites. In the Jamaican montane rain-forest the concentration of the tissue fluids of the parasitic Loranthaceae is in general higher than those of the host.[4] The same relationship has been found to obtain in desert Loranthaceae.[5]

As far as we are aware the relative electrolyte contents of the tissue fluids of parasite and host have not been determined heretofore.

In August, 1920, we had the opportunity while carrying out work for the U. S. Department of Agriculture at Sacaton, Arizona to measure the specific electrical conductivity, K, as well as the osmotic concentration in atmospheres, P, calculated from the freezing point lowering, Δ, of the expressed sap of the leaves of the host trees and of the stems of the leafless *Phoradendron californicum* parasitic on the leguminous trees *Acacia greggii* and

[1] MacDougal and Cannon, Pub. Carn. Inst. Wash., 1910, No. cxxix, P. 3, 25–49.

[2] MacDougal, *Bot. Gaz.*, 1911, lii, 249–260; *Bull. Torr. Bot. Club*, 1911, xxviii, 54–55, 473–480.

[3] Senn, *Verhandl. Naturf. Ges. in Basel*, 1913, xc, 179–183.

[4] Harris and Lawrence, *Amer. Jour. Bot.*, 1916, iii, 438–455.

[5] Harris, *Mem. Torr. Bot. Club*, 1918, xvii, 307–315.

Olneya tesota and of the leaves of the leafy *P. cockerellii* parasitic on *Populus wislizeni*, *Salix wrightii* and *Fraxinus attenuata*.

Sap was extracted after antecedent freezing of the tissues in an ice and salt mixture[1] to facilitate extraction[2] and the constants determined on the centrifuged sap.

Table I shows the average values of Δ and of P as determined from a published table.[3]

TABLE I.

FREEZING POINT LOWERING Δ, AND OSMOTIC CONCENTRATION, P.

Parasite and Host.	Δ			P		
	Para-site.	Host.	Differ-ence.	Para-site.	Host.	Differ-ence.
P. Californicum						
on *Acacia greggii*	2.81	2.21	+0.60	33.66	26.57	+7.09
on *Olneya tesota*	2.25	2.10	+0.15	26.98	25.24	+1.74
P. Cockerellii						
on *Populus wislizeni*............	1.92	1.84	+0.08	23.05	22.04	+1.01
on *Salix wrightii*..............	2.08	1.74	+0.34	24.98	20.88	+4.10
on *Fraxinus attenuata*..........	2.20	1.96	+0.24	26.47	23.47	+3.00

For each comparison the osmotic concentration of the tissue fluids of the parasite is higher than that of the host. Thus the results of earlier investigations in Jamaica and Arizona are confirmed.

TABLE II.

SPECIFIC ELECTRICAL CONDUCTIVITY, $K \times 10^5$, AND THE RATIO OF K TO Δ, $K/\Delta \times 10^5$.

Parasite and Host.	K			K/Δ		
	Para-site.	Host.	Differ-ence.	Para-site.	Host.	Differ-ence.
P. Californicum						
on *Acacia greggii*	2242	1509	+ 733	831	682	+149
on *Olneya tesota*..............	2471	2192	+ 279	1103	1052	+ 51
P. Cockerellii						
on *Populus wislizeni*..........	3061	1990	+1071	1598	1094	+504
on *Salix wrightii*..............	3101	1582	+1519	1488	908	+580
on *Fraxinus attenuata*.........	2399	1461	+ 938	1091	749	+342

The constants for specific electrical conductivity and for the ratio of electrical conductivity to freezing point lowering appear

---

[1] Gortner and Harris, *Pl. World*, 1914, xvii, 49–53.

[2] Dixon and Atkins, *Sci. Proc. Roy. Dublin Soc.*, 1913, N. S., xiii, 422–423; Gortner, Lawrence and Harris, *Biochem. Bull.*, 1916, v, 139–142.

[3] Harris and Gortner, *Amer. Jour. Bot.*, 1914, i, 75–78.

in Table II.   This shows that electrical conductivity, like freezing-point lowering, and the ratio K/Δ is higher in parasite than in host.

Thus it appears that there is some mechanism not as yet determined by which the mistletoe accumulates and retains in solution larger quantities of dissociated salts or organic acids than does the host.

It is possible that higher transpiration from the parasite might result in the accumulation in a purely mechanical manner of larger amounts of salts from the transpiration stream, but this is merely a suggestion requiring further investigation.

ON THE DIFFERENTIATION OF THE LEAF TISSUE FLUIDS OF LIGNEOUS AND HERBACEOUS PLANTS WITH RESPECT TO OSMOTIC CONCENTRATION AND ELECTRICAL CONDUCTIVITY.

By J. ARTHUR HARRIS, ROSS AIKEN GORTNER, AND JOHN V. LAWRENCE.

ON THE DIFFERENTIATION OF THE LEAF TISSUE FLUIDS OF LIGNEOUS AND HERBACEOUS PLANTS WITH RESPECT TO OSMOTIC CONCENTRATION AND ELEC- TRICAL CONDUCTIVITY.*

BY J. ARTHUR HARRIS, ROSS AIKEN GORTNER, AND JOHN V. LAWRENCE.

(*From the Department of Experimental Evolution and the Department of Botanical Research, the Carnegie Institution of Washington, Washington.*)

The existence of a differentiation of ligneous and herbaceous plants with respect to the magnitude of the osmotic concentration of the tissue fluids was first demonstrated in a strictly quantitative manner by work on the sap of the plants of the spring flora of the Arizona deserts[1] in the neighborhood of the Desert Botanical Laboratory, and on the terrestrial vegetation of the Jamaican montane rain forest.[2] These studies, in two geographically widely separated and climatically dissimilar regions, and an extensive series of unpublished observations demonstrate that the leaf tissue fluids of ligneous plants are characterized by an osmotic concentration materially higher than that of herbaceous forms.

The magnitude of the specific electrical conductivity, K, of the fluids must now be considered in comparison with osmotic concentration as measured by the freezing point lowering, Δ, for a series of plant species on which both of these constants were determined.

The determinations here considered were made on the north shore of Long Island during the spring and summer of 1914 and 1915. Leaf

---

* Studies carried out by the cooperation of the Department of Experimental Evolution and the Department of Botanical Research of the Carnegie Institution of Washington. The results will be published in full in the *Journal of Physical Chemistry*.
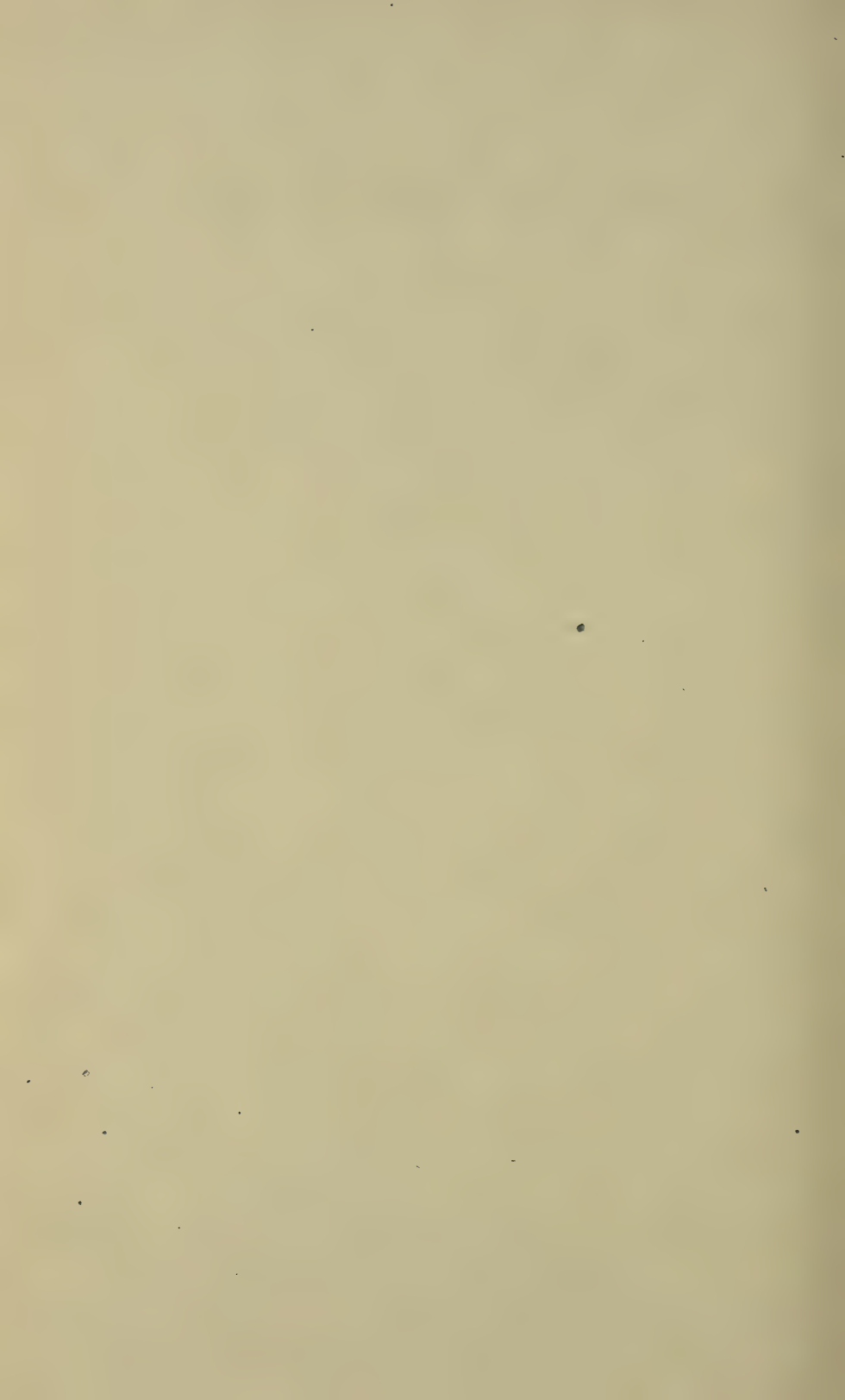
[1] Harris, J. A., Lawrence, J. V., and Gortner, R. A., *Phys. Researches*, 1916, ii, 1.

[2] Harris, J. A., and Lawrence, J. V., *Am. J. Bot.*, 1917, iv, 268.

THE OSMOTIC CONCENTRATION AND ELECTRICAL CONDUCTIVITY OF THE TISSUE FLUIDS OF LIGNEOUS AND HERBACEOUS PLANTS

BY J. ARTHUR HARRIS, ROSS AIKEN GORTNER AND JOHN V. LAWRENCE

THE OSMOTIC CONCENTRATION AND ELECTRICAL CONDUCTIVITY OF THE TISSUE FLUIDS OF LIGNEOUS AND HERBACEOUS PLANTS[1]

BY J. ARTHUR HARRIS, ROSS AIKEN GORTNER AND JOHN V. LAWRENCE

I. The Osmotic Concentration of the Leaf Tissue Fluids of Herbaceous and Ligneous Plants

An examination of earlier literature on the osmotic concentration of plant tissue fluids shows various suggestions of a difference in the osmotic concentration of the leaf fluids of herbaceous and ligneous plants. As early as 1911 Fitting[2] noted from his plasmolytic studies on desert plants that the lowest osmotic pressures are found in annuals and the highest in shrubs.

The existence of a differentiation of ligneous and herbaceous plants with respect to the magnitude of the osmotic concentration of their tissue fluids was first demonstrated in a strictly quantitative manner by work on the sap of the plants of the spring flora of the Arizona deserts[3] in the neighborhood of the Desert Botanical Laboratory. Because of the strongly contrasted environmental conditions in these southwestern deserts the growth forms are sharply differentiated. Following as closely as possible Thornber's classification of the growth forms,[4] thereby obviating any possible question of personal equation in the classification of the plants, we find the results for freezing point lowering given in Table I.[5]

[2] H. Fitting: Zeit. Bot., **3,** 209–275 (1911).

[3] J. Arthur Harris, J. V. Lawrence and R. A. Gortner: Phys. Res., **2,** 1–49 (1916).

[4] J. J. Thornber: Pub. Carnegie Inst. Wash., No. 113, pp. 103–112 (1909).

[5] Harris, Lawrence and Gortner: Loc. cit., pp. 45–46. The averages in the table are averages of species determinations, not of species means.

TABLE I

Average Freezing Point Lowering, Δ, of Different Growth Forms from Various Habitats in the Arizona Deserts

Growth form	Arroyo	Canyons	Rocky slopes	Bajada slopes	Salt spots	All habitats
Trees and shrubs	1.476	1.867	1.857	2.895	4.008	2.340
Dwarf and half shrubs	1.380	1.615	1.701	1.942	2.850	1.733
Perennial herbs	1.128	1.150	1.415	1.636	—	1.357
Winter annuals	1.075	1.094	1.269	1.759	1.960	1.227

TABLE II

Actual and Percentage Difference in Freezing Point Lowering, Δ, of Tissue Fluids of Ligneous and Herbaceous Plants of the Arizona Deserts

Growth form	Arroyo	Canyons	Rocky slopes	Bajada slopes	Salt spots	All habitats
All species	1.156	1.047	1.516	2.197	3.095	1.594
Ligneous species	1.441	1.751	1.753	2.523	3.776	2.055
Herbaceous species	1.080	1.117	1.325	1.708	1.960	1.261
Difference	0.361	0.634	0.428	0.815	1.816	0.794
Percentage difference	25.05	36.20	24.41	32.30	48.09	38.63

It is clear that the concentration of the leaf tissue fluids of trees and shrubs is higher than that of dwarf shrubs and half shrubs. The freezing point lowering found in the tissue fluids of both perennial herbs and winter annuals is, without exception, less than that observed in either of the ligneous growth forms.

Combining the growth forms into two groups, ligneous and herbaceous (Table II), the freezing point lowering of the leaf tissue fluids of the herbaceous plants is seen to be less in every instance than that of the ligneous species from the same habitats. The percentage differences, calculated by using the value of the ligneous plants as a base, range from 24.4 to 48.1.

Unfortunately, it is impossible to group the plant species of a temperate mesophytic region or of a tropical region with ample moisture and uniform temperature in exactly the same way as those from a subtropical desert of the type in which our first investigation was carried out. Biennials, and species growing as either biennials or annuals, which are so conspicuous a feature of a mesophytic temperate flora are practically wanting in the desert. The dwarf shrubs and half shrubs which are so dominant as an element in the desert flora studied are not abundant in the flora of the Eastern United States. The herbaceous annuals, prominent in both the xerophytic and mesophytic nearctic region, are but sparingly represented in moist subtropical and tropical regions.

A uniform classification of growth forms,is, therefore, impossible. Studies on the terrestrial vegetation of the Jamaican montane rain forest have shown the following results for the average of species averages of the freezing point depression[1] (Table III).

Without exception the values of Δ are higher for the ligneous than for the herbaceous species. The percentage differences, calculated as above, range from 22.1 to 30.3.

It is, therefore, fully demonstrated by these data from two geographically widely separated and climatically very

---

[1] J. Arthur Harris and J. V. Lawrence: Am. Jour. Bot., **4**, 287 (1917).

TABLE III

Actual and Percentage Difference in Freezing Point Lowering, Δ, of the Tissue Fluids of Ligneous and Herbaceous Plants of the Jamaican Montane Rain Forest

Growth form	Leeward slopes	Leeward ravines	Ridge forest	Windward habitats	All habitats
All species	1.007	0.822	0.914	0.743	0.881
Ligneous species	1.089	0.901	0.958	0.805	0.952
Herbaceous species	0.812	0.628	0.718	0.627	0.700
Difference	0.277	0.273	0.240	0.178	0.252
Percentage difference	25.43	30.29	25.05	22.11	26.47

dissimilar regions (and by extensive unpublished series) that the leaf tissue fluids of ligneous plants are characterized by an osmotic concentration materially higher than that of herbaceous forms.

The magnitude of the specific electrical conductivity of the fluids must next be considered in comparison with osmotic concentration as measured by the freezing point lowering.

II. The Osmotic Concentration and Specific Electrical Conductivity of the Leaf Tissue Fluids of Herbaceous and Ligneous Plants of a Mesophytic Flora

The determinations here considered were made on the north shore of Long Island during 1914 and 1915. Leaf tissue was collected in large test tubes. After freezing in an ice-salt mixture[1] to render the tissue permeable, as has been shown to be necessary by Dixon and Atkins[2] and by ourselves.[3] The sap was extracted as completely as possible by pressure, cleared by centrifuging and the freezing point lowering, Δ, was determined in the usual manner. Correction was made for the ice separating on undercooling by the formula

$$\Delta = 0.0125 \, u \, \Delta',$$

where u is the undercooling and Δ' the observed freezing point lowering in degrees.

The specific electrical conductivity, κ, of the sap was measured at 30° C in a Freas conductivity cell, standardized against $N/10$ KCl, which has a specific conductivity 0.01412 reciprocal ohms at 30°, by means of the ordinary meter bridge wire and resistance box of the physiological laboratory.

All determinations were made with as great care as possible, but there are many possible sources of error, and some selection of the constants to be used in the present paper seems

[1] R. A. Gortner and J. Arthur Harris: Plant World, **17**, 49–53 (1914).

[2] H. H. Dixon and W. R. G. Atkins: Proc. Roy. Soc. Dublin, **13**, 422–433 (1913). Also in Notes Bot. Sch. Trin. Coll., Dublin, **2**, 154–172 (1913).

[3] R. A. Gortner, J. V. Lawrence and J. Arthur Harris: Biochem. Bull., **5**, 139–142, pl. 1 (1916).

desirable. To avoid weighting the species upon which more than a single determination had been made, species averages are used whenever possible. These were determined as follows. The whole of the data which had been accumulated at various times during the two years were arranged together by species and all determinations which seemed obviously open to criticism were thrown out. The determinations for each species were then averaged and the deviation of each determination from the average for the species was calculated. All numbers which showed a deviation of more than ± 20 percent for either Δ, κ, or κ/Δ, were discarded, and a new average with deviations $< \pm 20$ percent determined from the remainder.

The inclusion of determinations differing from the average by as much as ± 20 percent might at first seem to represent great laxness of selection. One must remember, however, that these variations represent more than the errors of experimental measurement. They include all the differences due to seasonal and environmental influence as well as the errors of random sampling in the collection of the tissues. Thus the limits chosen probably represent rather stringent instead of lax selection. The detailed data are shown in Table IV.

Determining the usual statistical constants from the protocols of measurements we have the accompanying results (Table V) for the three growth forms, and for a combination of the two groups of ligneous plants.

The constants in Table V show that the mean freezing point lowering of the leaf tissue is greater, although perhaps not significantly greater in comparison with its probable error, in arborescent than in shrubby species. The tissue fluids of both trees and shrubs are characterized by a far greater freezing point lowering than those of herbaceous plants. The differences between trees and herbs, shrubs and herbs, and all ligneous plants and herbs are everal times as large as the probable error of the difference and hence unquestionably significant.

TABLE IV
Protocols of Determinations[1]
Constants for Trees

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Acer rubrum (2)	1.132	6585	7325
Aesculus Hippocastanum	0.722	10398	14401
Ailanthus glandulosa	1.480	15546	10504
Alnus rugosa (2)	1.368	9393	6863
Betula lenta	1.464	11080	7589
Betula lutea	1.221	10737	8833
Cynoxylon floridum (3)	1.119	12264	10957
Diospyros virginiana (2)	1.385	8586	6242
Gleditsia triacanthos (2)	1.250	11348	9083
Juglans cinerea	1.478	11802	8011
Padus virginiana (2)	1.754	8183	4659
Quercus coccinea	1.487	17287	11015
Quercus palustris	1.780	9871	5586
Quercus Prinus	1.655	12518	7568
Robinia Pseudacacia (3)	1.006	14364	14337
Salix alba vitellina	1.178	16513	14017
Tilia americana (2)	1.107	13814	12600
Vitis aestivalis	1.071	6107	5702
Vitis labrusca	0.892	6657	7463

Constants for Shrubs

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Amorpha fruticosa	1.104	10063	9115
Ampelopsis Veitchii	0.863	11460	13279
Aronia atropurpurea	1.165	10140	8703
Azalea nudiflora (3)	0.998	10696	10701
Benzoin aestivale (4)	1.104	11755	10744
Berberis vulgaris	1.555	9001	5788
Berberis vulgaris purpurea	1.598	8563	5358
Clethra alnifolia (3)	0.786	11075	14119
Comptonia peregrina	1.211	8018	6620
Cornus alternifolia	1.205	11886	9863
Epigaea repens	1.085	9526	8779
Gaylussacia frondosa	1.310	8777	6700
Hibiscus Syriacus (2)	1.120	17317	15499
Lonicera tatarica	1.644	11979	7286
Myrica carolinensis (2)	1.135	8619	7578

[1] Species are grouped primarily according to growth forms, as discussed in the paper. For convenience of reference they are alphabetically arranged under each growth form. The number in parentheses shows the number of individual determinations averaged to obtain the species constant. Those without numbers represent one determination only.

<div align="center">

TABLE IV (*Continued*)
Constants for Shrubs (*continued*)

</div>

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Parthenocissus quinquefolia	0.943	8426	8935
Prunus sp. (2)	1.514	18084	12104
Rhus glabra (3)	1.286	12131	9475
Rosa virginiana	1.043	10190	9769
Rubus argutus (2)	1.199	11456	9483
Rubus hispidus	0.818	11841	14475
Sambucus canadensis (6)	1.065	15583	14670
Smilax rotundifolia	1.237	11466	9269
Solanum Dulcamara (4)	0.914	17327	19146
Sorbaria sorbifolia (2)	1.297	16119	12453
Toxicodendron Toxicodendron	1.135	10190	8977
Toxicodendron Vernix	1.391	9989	7181
Uva-ursi Uva-ursi	1.218	5856	4809
Vaccinium angustifolium	0.965	8449	8755
Vaccinium atlanticum	1.366	5896	4316
Vaccinium corymbosum	1.561	8038	5149
Vaccinium vacillans	0.948	6768	7139
Viburnum acerifolium (3)	1.057	13153	12429
Viburnum cassinoides	1.300	7467	5743
Viburnum dentatum (2)	1.272	11052	8899
Xolisma ligustrina (2)	0.977	9355	9730

<div align="center">Constants for Herbs</div>

Achillea lanulosa	0.815	16645	20423
Achillea Millefolium	0.738	13622	18457
Agrimonia gryposepala (3)	0.861	14508	16899
Agrimonia sp. (2)	0.875	13726	15726
Alsine media (4)	0.605	16388	27185
Ambrosia elatior	0.807	22250	27571
Ambrosia trifida	0.754	20072	26620
Anaphalis margaritacea (2)	0.940	19785	21140
Antenarria plantaginifolia	0.743	16987	22862
Anthemis Cotula	0.604	15671	25945
Aralia nudicaulis	1.386	9515	6865
Asclepias pulchra	0.593	13703	23116
Asparagus officinalis	1.545	18795	12165
Aster macrophyllus (2)	0.646	17075	27124
Aureolaria Pedicularia	1.411	19195	13603
Aureolaria villosa	1.174	16179	13781
Baptisia tinctoria (3)	1.016	7832	7752
Barbarea Barbarea (4)	0.785	17133	21983
Barbarea stricta (2)	0.795	16038	20181
Brassica juncea (2)	0.728	15836	21787

TABLE IV (*Continued*)
Constants for Herbs (*continued*)

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Brassica napus	0.803	17897	22287
Brassica nigra	0.789	19257	24407
Cardamine pennsylvanica (2)	0.670	18141	27083
Carex scoparia	1.017	16447	16172
Chelidonium majus (4)	1.000	11345	11347
Chenopodium album	0.991	24054	24272
Chenopodium sp.	0.936	26845	28680
Chimaphila maculata	0.984	6349	6452
Chrysanthemum Leucanthemum (2)	0.967	15770	16305
Chrysopsis mariana (2)	0.806	14850	18447
Cichorium Intybus (3)	0.796	18711	23496
Cimicifuga racemosa	0.937	15235	16259
Circaea lutetiana (4)	0.489	10397	21363
Commelina communis	0.422	12176	28861
Convallaria majalis	0.829	12574	15167
Convulvulus arvensis	0.937	18272	19500
Crocanthemum canadense	0.741	8732	11784
Daucus Carota (2)	1.143	20394	17836
Deringa canadensis (2)	0.934	21182	22810
Dianthus Armeria	1.009	19038	18868
Erechtites hieracifolia	0.506	12422	24549
Erigeron annuus (3)	0.808	14070	17427
Erigeron ramosus	1.153	10355	8981
Eupatorium perfoliatum	0.583	13457	23082
Eupatorium trifoliatum	0.813	20372	25057
Euthamia graminifolia (2)	0.936	14846	15926
Euthamia tenuifolia	0.721	14936	20730
Fagopyrum Fagopyrum	0.540	13198	24440
Fragaria vesca americana	1.098	9384	8546
Fragaria virginiana	0.998	12828	12853
Galinsoga parviflora (2)	0.601	16392	27329
Galium Aparine (5)	0.722	12486	17365
Geranium maculatum (5)	0.768	9922	13000
Geranium pusillum	0.789	17400	22053
Geranium rotundifolium	0.888	17829	20077
Geum canadense (2)	1.192	19014	16128
Gratiola aurea	0.618	12518	20255
Hemerocallis fulva	0.940	9459	10062
Hieracium sp.	0.783	15990	20421
Hypericum mutilum	0.865	11599	13409
Hypericum perforatum (3)	1.002	12181	12239
Hypericum punctatum	0.833	12473	14973
Impatiens biflora (5)	0.518	11784	22857
Ionactis linariifolius	0.881	13566	15398

TABLE IV (*Continued*)

Constants for Herbs (*continued*)

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Lactuca virosa	0.681	18264	26819
Lappula virginiana	0.644	15546	24139
Lathyrus latifolius (2)	0.942	10562	11219
Leontodon Taraxacum	0.707	13622	19267
Leonurus Cardiaca (4)	0.903	18654	20769
Leptilon canadense (2)	0.810	14509	17957
Lespedeza capitata	0.946	12323	13026
Lespedeza frutescens	1.035	7315	7067
Lespedeza hirta (2)	0.794	9734	12271
Lespedeza violacea	0.984	10834	11010
Lespedeza virginica	0.836	10566	12638
Linaria canadensis (3)	0.580	10495	18167
Linaria Linaria (2)	0.847	10232	12088
Lychnis alba	0.793	20680	26078
Lychnis dioica	0.711	18126	25493
Lycopodium obscurum	0.874	7928	9070
Lycopus sessilifolius	0.625	16120	25792
Lycopus virginicus (2)	0.538	13554	25234
Lysimachia Nummularia (2)	0.747	12907	17397
Lysimachia quadrifolia (3)	0.634	11070	17660
Lysimachia terrestris	0.717	9076	12658
Medeola virginiana (2)	0.833	13543	16371
Medicago lupulina	1.068	12371	11583
Melampyrum lineare (2)	1.164	15427	13248
Melilotus alba	1.119	10358	9256
Mentha citrata	0.751	13352	17807
Monarda didyma (2)	0.694	13636	19649
Monarda fistulosa	1.037	12574	12125
Nepeta Cataria	0.724	15740	21740
Nepeta hederacea (2)	0.650	12805	19743
Oenothera muricata	0.711	11284	15885
Oenothera Oakesiana	0.726	12181	16778
Ornithogalum umbellatum	0.713	9268	12998
Osmunda regalis	1.180	15925	13495
Panicum clandestinum	0.764	14877	19535
Persicaria Hydropiper	0.707	12225	17291
Persicaria Persicaria	0.607	11321	18650
Persicaria punctata	0.586	12473	21285
Phlox paniculata	0.737	14755	20020
Physalis heterophylla	0.704	15175	21555
Phytolacca decandra	0.726	14063	19370
Plantago lanceolata	0.867	13622	15711
Plantago media	0.775	18500	23870

<div align="center">

TABLE IV (*Continued*)

Constants for Herbs (*continued*)

</div>

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Plantago Rugelii	0.789	18192	23057
Polygonatum commutatum	1.014	10920	10769
Polygonum aviculara	0.593	11373	19178
Portulaca oleracea (2)	0.598	15962	26713
Potentilla canadensis	0.935	12225	13074
Potentilla monspeliensis	1.050	18205	17338
Prunella vulgaris	0.638	9740	15266
Pteridium aquilinum	1.555	15671	10077
Ranunculus abortivus	1.231	10876	8835
Ranunculus bulbosus (2)	1.016	12397	12203
Ranunculus recurvatus (2)	0.998	10574	10690
Ranunculus sceleratus	0.941	17264	18346
Ranunculus septentrionalis	1.026	13677	13330
Rudbeckia hirta (2)	0.863	16616	19311
Rumex Acetosella (2)	0.531	8273	15596
Rumex crispus (2)	0.657	17085	25976
Rumex hastatulus	0.563	11883	21106
Rumex obtusifolius (3)	0.706	15227	21602
Saponaria officinalis	0.970	9780	10082
Scirpus polyphyllus	0.894	16968	18991
Scrophularia leporella (2)	0.834	14553	17541
Sedum purpureum (2)	0.471	4061	8707
Sericocarpus asteroides (4)	0.703	12302	17549
Silene latifolia	0.872	21469	24620
Sinapis arvensis	0.888	18264	20567
Sisymbrium Nasturtium-aquaticum (2)	0.652	17274	26513
Sisymbrium sp.	0.803	18960	23611
Solidago altissima	0.959	15423	16090
Sol dago bicolor	0.815	14231	17461
Solidago juncea (3)	1.074	14710	13781
Solidago odora	0.911	15546	17064
Solidago rugosa	1.126	13037	11580
Solidago sp.	0.984	15483	15734
Spathyema foetida (2)	1.039	16424	15871
Specularia perfoliata	0.792	14756	18631
Tanacetum vulgare (2)	0.972	16164	16645
Thalictrum dioicum	0.988	14233	14406
Tovara virginiana (2)	0.451	9703	21991
Tridentalis americana	0.844	12471	14776
Trientalis borealis (2)	0.899	12194	13679
Trifolium agrarium	0.949	8323	8770
Trillium cernuum (2)	1.022	18054	17748
Unifolium canadense (4)	1.002	11655	11650

TABLE IV (*Continued*)

Constants for Herbs (*continued*)

	Δ	$\kappa \times 10^6$	$\kappa/\Delta \times 10^6$
Urtica gracilis	1.175	17213	14656
Uvularia perfoliata	0.833	12372	14852
Vagnera racemosa (8)	1.041	12313	11880
Veratrum viride (2)	0.845	14073	16751
Verbascum Blattaria	0.812	13512	16640
Verbena urticifolia (5)	0.831	12651	15342
Veronica officinalis (2)	0.929	13730	14767
Vicia Cracca	0.875	12225	13971
V ola cucullata (2)	0.715	10030	14234
Viola pa'lens	0.763	12225	16022
Washingtonia longistylis	1.123	19195	17092
Xanthoxalis corniculata	0.796	21469	26971

Expressing the differences in percentages of the constants for ligneous forms, we note that the value for trees and shrubs is 30.46 percent higher than that of herbaceous plants.

These results are, therefore, in excellent agreement with those found in the Arizona deserts and in the Jamaican rain forest.

The constants set forth in Table VI show that the specific electrical conductivity for shrubs is slightly lower than that for trees. The difference is, however, smaller than its probable error. The differences between the conductivities of the leaf tissue fluids of trees and herbs, shrubs and herbs, and all ligneous species and herbs, are several times as large as their probable errors and show that the conductivity is distinctly higher in herbaceous than in ligneous plants.

The constants for the ratio of electrical conductivity to freezing point lowering, κ/Δ, appear in Table VII.

The entries in this table show that the ratio of conductivity to freezing point lowering is lower in trees than in shrubs, although the difference cannot be considered significant in comparison with its probable error. The ratio of conductivity to freezing point depression is much smaller in both trees and shrubs than it is in herbs. The ratios ($\times 10^6$) are

TABLE V

Statistical Constants for Freezing Point Lowering, Δ

	Mean	Standard deviation	Coefficient of variation
(1) Trees, N = 19	1.2921 ± 0.0429	0.2772 ± 0.0303	21.46 ± 2.45
(2) Shrubs, N = 36	1.1775 ± 0.0244	0.2174 ± 0.0173	18.46 ± 1.52
Difference (1) — (2)	+0.1146 ± 0.0493	+0.0598 ± 0.0348	+3.00 ± 2.88
(1 + 2) Trees and shrubs, N = 55	1.2171 ± 0.0224	0.2459 ± 0.0158	20.20 ± 1.35
(3) Herbs, N = 162	0.8464 ± 0.0105	0.1986 ± 0.0074	23.46 ± 0.93
Difference (1) — (3)	+0.4457 ± 0.0441	+0.0786 ± 0.0311	−2.00 ± 2.62
Difference (2) — (3)	+0.3311 ± 0.0265	+0.0188 ± 0.0188	−5.00 ± 1.78
Difference (1 + 2) — (3)	+0.3707 ± 0.0247	+0.0473 ± 0.0174	−3.26 ± 1.64

TABLE VI

Statistical Constants for Specific Electrical Conductivity, $\kappa \times 10^6$

	Mean	Standard deviation	Coefficient of variation
(1) Trees, N = 19	11213 ± 494	3195 ± 350	28.49 ± 3.36
(2) Shrubs, N = 36	10770 ± 339	3019 ± 240	28.03 ± 2.40
Difference (1) — (2)	+443 ± 599	+176 ± 425	+0.46 ± 4.13
(1 + 2) Trees and shrubs, N = 55	10923 ± 281	3088 ± 199	28.27 ± 1.96
(3) Herbs, N = 162	14308 ± 192	3624 ± 136	25.33 ± 1.01
Difference (1) — (3)	−3095 ± 530	−429 ± 375	+3.16 ± 3.51
Difference (2) — (3)	−3538 ± 389	−605 ± 275	+2.70 ± 2.60
Difference (1 + 2) — (3)	−3385 ± 340	−536 ± 240	+2.94 ± 2.20

Table VII

Statistical Constants for Ratio of Specific Electrical Conductivity to Freezing Point Lowering, $\kappa/\Delta \times 10^6$

	Mean	Standard deviation	Coefficient of variation
(1) Trees, N = 19	9092 ± 462	2987 ± 327	32.85 ± 3.96
(2) Shrubs, N = 36	9529 ± 372	3312 ± 263	34.75 ± 3.07
Difference (1) — (2)	—437 ± 593	—325 ± 419	—1.90 ± 5.01
(1 + 2) Trees and shrubs, N = 55	9378 ± 292	3210 ± 206	34.23 ± 2.44
(3) Herbs, N = 162	17674 ± 282	5326 ± 200	30.13 ± 1.22
Difference (1) — (3)	—8582 ± 541	—2339 ± 383	+2.72 ± 4.14
Difference (2) — (3)	—8145 ± 466	—2014 ± 330	+4.62 ± 3.31
Difference (1 + 2) — (3)	—8296 ± 405	—2116 ± 287	+4.10 ± 2.73

9092 : 17674 in the case of trees and herbs and 9529 : 17674 in the case of shrubs and herbs. Since the ratio does not differ significantly in trees and shrubs it is quite proper to combine them. The average value of $\kappa/\Delta \times 10^6$ in all ligneous plants is 9378 ± 292 as compared with 17674 ± 282 in herbs. Thus the ratio κ/Δ is about 90 percent higher in herbaceous than in ligneous plants.

It seems desirable from the physiological side to determine whether conductivity or osmotic concentration is more nearly a constant for the species of a region. Furthermore, it is of interest to determine whether either of the three physico-chemical constants considered is significantly more variable in one class of plants (trees, shrubs, or herbs) than another.

Comparison of the variability of electrical conductivity and of osmotic pressure can only be made by means of the relative variation constants. The differences in the coefficients of variation of κ and Δ for the three growth forms are given in Table VIII.

<div align="center">

TABLE VIII

Difference in Coefficients of Variation of Electrical Conductivity and Freezing Point Lowering

</div>

Growth form	Difference C. V.$_\kappa$ — C. V.$_\Delta$
Trees	7.03 ± 4.16
Shrubs	9.57 ± 2.84
Trees and Shrubs	8.07 ± 2.38
Herbs	1.87 ± 1.36

The comparison shows that for each of the growth forms investigated the variability of electrical conductivity from species to species is greater than that of osmotic concentration. The differences are conspicuously greater in ligneous plants (in which the conductivity is on the average small as compared with the osmotic concentration) than in herbaceous species in which the electrical conductivity is both absolutely and relatively much larger. As a matter of fact, the varia-

bility of κ and of Δ cannot be asserted to be significantly different in herbaceous species.

Turning to the question of the relative variability of the physico-chemical constants in the different growth forms we note that the variability of the freezing point lowering of trees is numerically greater than that of shrubs as measured by both standard deviation and coefficient of variation. The difference in the variability of the two groups of ligneous forms is not, however, sufficiently large to be considered significant in comparison with its probable error. The standard deviation of Δ in herbaceous forms is lower than that in either trees or shrubs. The average osmotic concentrations, as measured by Δ, is also lower in these forms. In consequence the relative variability as measured by the coefficient of variation is higher in herbaceous species.

The variability from species to species of the electrical conductivity does not differ, as far as the data now available show, in the two groups of ligneous plants. The standard deviations of the conductivities of herbaceous plants are higher than those of ligneous species, but since the average conductivities are also higher, their relative variabilities as measured by the coefficient of variation are somewhat lower.

Finally, the results of the ratio κ/Δ show that the standard deviations are far higher in herbaceous than in ligneous plants. The mean value of the ratio is also far higher, and as a result the coefficients of variation of ligneous species are higher.

Taking the results for variability as a whole, they seem to indicate that there is little difference between ligneous and herbaceous forms.

The foregoing results show clearly that the osmotic concentration is higher while the electrical conductivity is lower in the tissue fluids of ligneous than in those of herbaceous species.

The results for freezing point lowering are drawn from three climatically highly dissimilar regions. While those for electrical conductivity are based on determinations from one region only, they represent two years work.

TABLE IX

Average Value of $\kappa \times 10^6$ and of $\kappa/\Delta \times 10^6$ for Species Classified according to Osmotic Concentration

Δ	Mean value of $\kappa \times 10^6$				Mean value of $\kappa/\Delta \times 10^6$			
	Trees	Shrubs	Trees and shrubs	Herbs	Trees	Shrubs	Trees and shrubs	Herbs
0.465	—	—	—	9752	—	—	—	21094
0.558	—	—	—	12940	—	—	—	22619
0.651	—	—	—	14785	—	—	—	22990
0.744	10398	11075	10737	14234	14401	14119	14260	19237
0.837	—	11651	11651	14841	—	13877	13877	17975
0.930	6657	10243	9525	15506	7463	10994	10288	16689
1.023	14364	11795	12224	13352	14337	11460	11939	13190
1.116	9693	11245	10624	13919	9146	10115	9728	12439
1.209	12866	9804	10824	15772	10644	8125	8964	13357
1.302	—	11109	11109	—	—	8654	8654	—
1.395	8990	7943	8466	14355	6553	5749	6151	10234
1.488	13929	18084	14760	—	9280	12104	9845	—
1.581	—	8534	8534	17233	—	5432	5432	11121
1.674	12518	11979	12249	—	7568	7286	7427	—
1.767	9027	—	9027	—	5123	—	5123	—

TABLE X

Average Value of Δ and $\kappa/\Delta \times 10^6$ for Species Classified according to Specific Electrical Conductivity

$\kappa \times 10^6$	Mean value of Δ				Mean value of $\kappa/\Delta \times 10^6$			
	Trees	Shrubs	Trees and shrubs	Herbs	Trees	Shrubs	Trees and shrubs	Herbs
4000	—	—	—	0.471	—	—	—	8707
5000	—	—	—	—	—	—	—	—
6000	1.071	1.292	1.218	0.984	5702	4563	4942	6452
7000	1.012	1.124	1.068	1.035	7394	6441	6918	7067
8000	1.754	1.170	1.286	0.842	4659	7365	6824	10297
9000	1.377	1.315	1.332	0.841	6553	7031	6894	11210
10000	1.251	1.154	1.178	0.825	9994	8754	9064	13630
11000	1.312	1.059	1.143	0.868	8502	10958	10139	13444
12000	1.299	1.211	1.236	0.798	9484	10369	10116	16303
13000	1.655	1.057	1.356	0.791	7568	12429	9999	17254
14000	1.056	—	1.056	0.804	13469	—	13469	17593
15000	—	—	—	0.861	—	—	—	17747
16000	1.480	1.181	1.280	0.862	10504	13562	12542	20036
17000	1.333	1.017	1.174	0.814	12516	17323	14919	21658
18000	—	1.514	1.514	0.837	—	12104	12104	22167
19000	—	—	—	1.063	—	—	—	18904
20000	—	—	—	0.912	—	—	—	22663
21000	—	—	—	0.848	—	—	—	25120
22000	—	—	—	0.807	—	—	—	27571
23000	—	—	—	—	—	—	—	—
24000	—	—	—	0.991	—	—	—	24272
25000	—	—	—	—	—	—	—	—
26000	—	—	—	—	—	—	—	—
27000	—	—	—	0.936	—	—	—	28680

While the materials are rather too meagre for exhaustive statistical analysis they have been classified according to the magnitude of Δ and the average value of κ and of κ/Δ for each class of Δ has been determined.   The results are presented in Tables IX and X.[1]

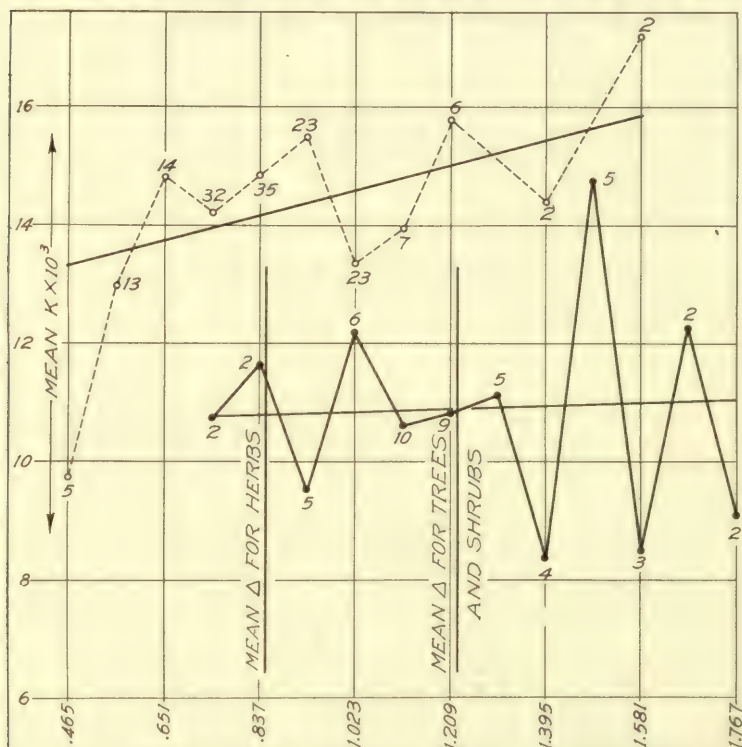Diagram 1, which shows by the position of the two means

Diagram 1

Average values of specific electrical conductivity, κ, of leaf tissue fluids of ligneous and herbaceous species classified according to freezing point depression, Δ.

on the scale for Δ the conspicuous differentiation of ligneous and herbaceous plants for this constant, also brings out clearly the fact that for each grade of Δ the ligneous forms have a lower electrical conductivity than the herbaceous species.   The mean values for κ for the two growth forms lie

[1] The class interval of Δ has been selected to represent 5 percent of the molecular lowering taken as $\Delta = 1.86$.

at the points where the straight lines fitted to the means intersect the verticals representing the mean values of the freezing point lowering.

Diagram 2

Average osmotic concentration, in terms of freezing point lowering, of leaf tissue fluids of ligneous and herbaceous species classified according to specific electrical conductivity, $\kappa \times 10^6$.

The mean values of Δ for various grades of κ are shown in Table X and represented graphically in Diagram 2. The differentiation of ligneous and herbaceous forms with respect

For trees:

$$r = +0.127 \pm 0.152, \kappa \times 10^6 = 9322 + 1464 \Delta$$
$$\Delta = 1.1685 + 0.00001102 \kappa \times 10^6$$

For shrubs:

$$r = -0.079 \pm 0.112, \kappa \times 10^6 = 12059 - 1095 \Delta$$
$$\Delta = 1.2385 - 0.00000567 \kappa \times 10^6$$

For trees and shrubs:

$$r = +0.022 \pm 0.091, \kappa \times 10^6 = 10591 + 273 \Delta$$
$$\Delta = 1.1985 + 0.00000153 \kappa \times 10^6$$

For herbs:

$$r = +0.150 \pm 0.052, \kappa \times 10^6 = 11997 + 2730 \Delta$$
$$\Delta = 0.7293 + 0.00000819 \kappa \times 10^6$$

The correlations between the freezing point lowering and the electrical conductivity of the sap of ligneous plants are of a very low order and statistically insignificant in comparison with their probable errors. The value for shrubs is actually negative in sign. That for trees and shrubs together is sensibly zero. The coefficient for herbaceous plants is also low, but may indicate a slight relationship between the two constants, higher values of Δ being associated with higher values of κ and *vice versa*.

These results show that, in the vegetation of the glacial moraines of Long Island at least, there is practically no relationship between the concentration of ionized electrolytes and of total solutes (molecules and ions) in the leaf tissue fluids.

If there be little relationship between the magnitudes of κ and Δ one might expect the value of the ratio κ/Δ to decrease as Δ becomes larger and to increase as κ increases in magnitude. This is shown in Diagrams 3 and 4 to be the case.

The explanation of the lower conductivity of the tissue fluids of ligneous plants presents a problem for future investigation.

It may be suggested that the arborescent forms are exposed to greater insolation, and in consequence show greater photosynthetic activity. A fragment of evidence in this

direction is furnished by the fact that Δ increases while the value κ/Δ decreases with height in arborescent plants.[1]
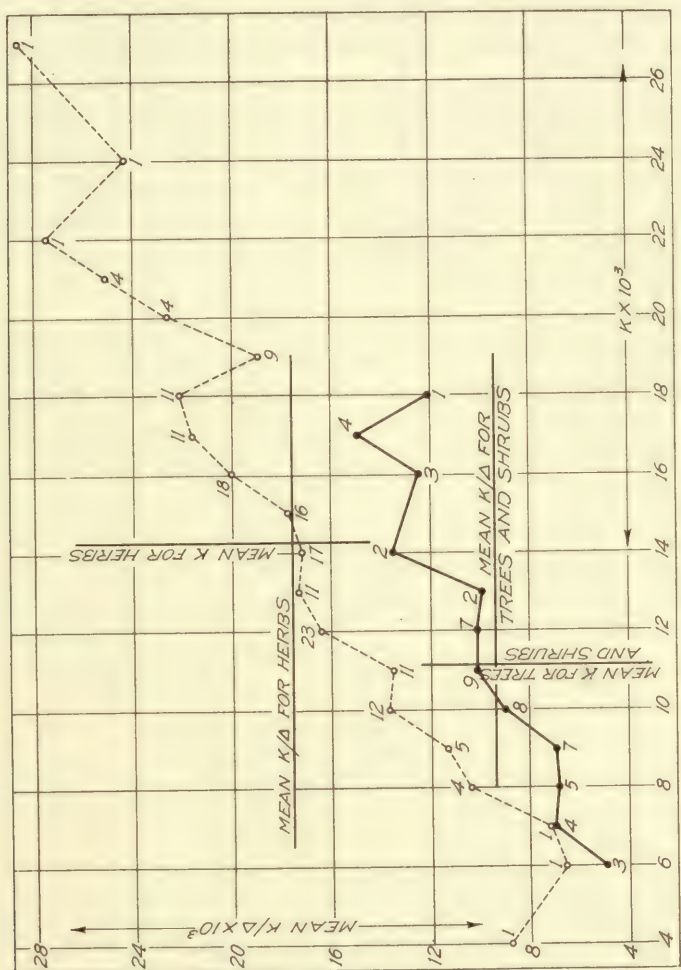

Diagram 4

Mean ratio of electrical conductivity to freezing point depression, κ/Δ, in leaf tissue fluids of ligneous and herbaceous species classified according to electrical conductivity, κ.

Summary

Studies in the Arizona deserts, in the Jamaican montane

[1] J. Arthur Harris, R. A. Gortner and John V. Lawrence: Bull. Torr. Bot. Club, **44**, 267–286 (1917).

rain forest, and in the mesophytic habitats of the north shore of Long Island, have shown that the osmotic concentration, as measured by the cryoscopic method, is far higher in the leaf tissue fluids of ligneous than of herbaceous species.

A large series of determinations in the various non-halophytic habitats of the north shore of Long Island shows that the specific electrical conductivity of the expressed leaf tissue fluids of ligneous species is lower than that of herbaceous species. This shows that while the concentration of ionized electrolytes is lower in ligneous than in herbaceous forms, the reverse is true for total solutes.

Because of the wide geographic range and the great diversity of conditions (xerophytic, mesophytic and hygrophytic) under which the investigations on osmotic concentration were carried out, there can be no reasonable doubt but that the differentiation of ligneous and herbaceous plants with respect to the magnitude of their osmotic concentration (Δ) is a general biological law. Until confirmed by investigations in other regions presenting different conditions for plant growth[1] the results for conductivity cannot be asserted to be of universal validity.

The Carnegie Institution
of Washington

---

[1] These investigations are now in progress.

50 (1632)
Maximum values of osmotic concentration in plant tissue fluids.

By J. ARTHUR HARRIS, R. A. GORTNER, W. F. HOFMANN, and A. T. VALENTINE.

[*From the Carnegie Station for Experimental Evolution, Cold Spring Harbor, N. Y., and the University of Minnesota, St. Paul, Minnesota.*]

The observations of a number of botanists have shown that extremely high concentrations may characterize plant tissue fluids, especially when the plants[1] occur in a highly saline substratum. To Fitting[2] belongs the credit of first demonstrating that extremely high osmotic concentrations are found in some desert plants,[3] although Drabble and Lake[4] and Drabble and Drabble[5] had preceded him in showing the fundamental relationship between environmental conditions and the osmotic concentration of plant tissue fluids.

As early as 1902 Cavara reported cryoscopic determinations on saps of high concentration[6] and in 1905 gave results in full[7] for a large series of determinations made at Cagliari. His maximum values were found in the sap of halophytes growing in localities where the concentration of the soil solution progressed with the advance of the season. He reports freezing point depressions of 7.25° in *Obione portulacoides*, 7.48° in *Salicornia fruticosa*, and 7.25° to 8.50° in *Halocnemum strobilaceum*. His determinations were, however, made on sap extracted without the antecedent treatment necessary to render the tissues permeable as has been

---

[1] Our present observations apply to the tissue fluids of flowering plants only. No attempt is made here to discuss the concentrations found in such lower organisms as those studied by G. J. Peirce (Pub. Carn. Inst. Wash., 1914, No. 193, p. 47–69) or G. Senn (*Verh. Schw. Naturf. Ges.*, 1911, xciv).

[2] H. Fitting, *Zeitschr. f. Bot.*, 1911, iii, 209–275.

[3] Fitting found a number of species of plants in the North-African deserts, the leaf cells of which were not plasmolyzed by 3 gram molecular KNO_3 solution. Theoretically potassium nitrate of this concentration should be the equivalent of about 100 atmospheres. The technical difficulties of applying the plasmolytic method are such as to lead one to question its value as a means of investigating in a quantitative manner the unusually high concentrations found in desert plants.

[4] E. Drabble and H. Lake, *New Phytologist*, 1905, viii, 189.

[5] E. Drabble and H. Drabble, *Biochem. Jour.*, 1907, ii, 117.

[6] F. Cavara, *Rendic. Congr. Bot. Palermo*, 1902, 66.

[7] F. Cavara, *Contrib. Biol. Veg.*, 1905, iv, 41–84.

shown to be necessary by Dixon and Atkins[8] and others.[9]   His constants are, therefore, as pointed out by Atkins,[10] probably sub-maximum because of incomplete extraction.

Work on the spring flora of the Arizona deserts[11] was probably carried out in a manner to obviate the objections to the preceding studies.   In this series the maximum concentrations were found in *Atriplex canescens*, a shrub of the salt spots, in which $\Delta = 5.65$, $P = 67.5$, and in *Mortonia scabrella*, a small shrub of the mesa-like slopes, for which one determination gave $\Delta = 4.78$, $P = 57.2$.

Concentrations of about fifty atmospheres have been demonstrated in the leaf tissue fluids of more or less sclerophyllous trees *Capparis ferruginea* and *Guaiacum officinale* and in those of the succulent-leaved halophytic half shrub *Batis maritima* of the saline coastal flats of Jamaica.[12]   Cryoscopic studies on mangrove vegetation[13] have indicated maximum concentrations of about fifty atmospheres in *Avicennia nitida*, although two questionable determinations indicated seventy atmospheres.   Using plasmolytic methods, von Faber[14] reports concentrations ranging from 24 to 72 atmospheres in East Indian species of the mangrove association.

During the summer of 1920, while engaged in field operations in collaboration with the Department of Agriculture in the Great Salt Lake region, we had the opportunity of making several hundred determinations of the osmotic concentration of plant tissue fluids by the cryoscopic method.   These measurements were made on sap extracted after freezing of the tissues[15] and with such care as to render the results reasonably free from criticism.   Such a series, based on species which have for ages been subject to the influence of the highly saline substratum afforded by the

[8] H. H. Dixon and W. R. G. Atkins, *Proc. Roy. Dublin Soc.*, 1913, N. S., xiii, 422–433.

[9] R. A. Gortner, J. V. Lawrence, and J. Arthur Harris, *Biochem. Bull.*, 1916, v, 139–142, *pl. 1*.

[10] W. R. G. Atkins, "Some Recent Researches in Plant Physiology," London, 1916, 94.

[11] J. Arthur Harris, J. V. Lawrence, and R. A. Gortner, *Phys. Res.*, 1916, ii, 1–49.

[12] J. Arthur Harris and J. V. Lawrence, *Bot. Gaz.*, 1917, lxiv, 285–305.

[13] J. Arthur Harris and J. V. Lawrence, *Biol. Bull.*, 1917, xxxii, 202–211.

[14] F. C. von Faber, *Ber. Deutch. Bot. Ges.*, 1913, xxxi, 277–281.

[15] R. A. Gortner and J. Arthur Harris, *Pl. World*, 1914, xvii, 49–53.

bed of the ancient Lake Bonneville, should furnish some indication of the maximum concentration[16] to be found in the leaf tissue fluids of flowering plants.

While high concentrations were demonstrated in a number of species, the highest was found in the typical salt desert half-shrub *Atriplex confertifolia*. It alone will be considered.

Two collections made on the rocky cliffs of Stansbury Island, Great Salt Lake, on July 14 gave freezing point depressions of 6.96° and 7.97°. If we may use the formula of Lewis,[17] upon which published tables of osmotic concentration have been based,[18] these depressions indicate osmotic pressures of 82.9 and 94.7 atmospheres respectively.

The highest concentrations were found in plants growing on the low ridges in the salt-flats[19] along the southern shore of Great Salt Lake. A determination on material collected July 16 gave $\Delta = 6.22$, $P = 74.2$.

On July 18 a determination on plants in about the same type of locality gave $\Delta = 10.00$, $P = 118.5$. Finally, on July 27 a determination made in this locality on the leaves of this species indicated a freezing point lowering of 13.0°. The equation used would indicate a concentration of 153.1 atmospheres.[20]

These determinations show that concentrations measured by a depression of 13.0°, presumably the equivalent of 153 atmospheres, maybe found in the tissue fluids of apparently normal leaves.

---

[16] A difficulty in work on the leaves of desert plants lies in the fact that the maximum concentrations must be expected during the periods of more extreme drought. During such periods the saps may become concentrated by desiccation merely. We know very little concerning the functional activities of such leaves or whether they are retained after the beginning of a period of more adequate moisture. There is, therefore, the possibility that leaves which are too desiccated to be longer functional may be utilized for determinations and indicate concentrations which are really larger than those in which the metabolic processes of the cells may be normally carried on. We believe that except as specifically indicated, the concentrations here recorded were determined on leaves in fairly normal condition.

[17] G. N. Lewis, *Journ. Amer. Chem. Soc.*, 1908, xxx, 668–683.

[18] J. Arthur Harris and R. A. Gortner, *Amer. Jour. Bot.*, 1914, i, 75–78; Harris, *Amer. Jour. Bot.*, 1915, ii, 418–419.

[19] T. H. Kearney, L. J. Briggs, H. L. Shantz, J. W. McLane, and R. L. Piemisel, *Jour. Agr. Res.*, 1914, i, 365–417, pl. 52–58.

[20] A sample from *Atriplex nuttallii* showed a freezing point lowering of about 14.4°, indicating a concentration of 169.3 atmospheres. The leaves appeared more dried than those of *Atriplex confertifolia*, and we are inclined to await further measurements before accepting this constant.

75 (1450)

The transformation of the plant ovule into an ovary.

By J. Arthur Harris.

[*From the Station for Experimental Evolution, Carnegie Institution of Washington, Cold Spring Harbor, N. Y.*]

In plants there is a rather wide capacity for the development of organs of various kinds from primordia normally destined to produce quite different structures. For example, leaves may replace petals, stamens or carpels; petals may occur in the place of stamens or carpels. The transformation of stamens into carpels is a well-known phenomenon.

Furthermore, the continued development of a growing point the activity of which is usually terminated by the formation of some highly specialized organ, such as the flower or fruit, is quite familiar to those concerned with problems of variation.

Among these morphological abnormalities the continued meristematic activity of the axis which is normally terminated by the formation of the ovary is of very rare occurrence. It is, however, regularly found, although in a small and variable percentage of the cases, in one of the passion flowers, *Passiflora gracilis*. Here prolification of the fruit consists in the formation of series of carpels, which may or may not be ovuliferous, within the normal fruit. The mass of accessory carpels thus formed may be so large as to rupture the fruit wall.

While the occurrence of prolification may be regarded as a heritable characteristic in *P. gracilis* the abnormality is of relatively rare occurrence. Physico-chemical factors must, therefore, determine the occurrence of prolification in certain fruits and its absence from others.[1]

---

[1] A prolonged effort to demonstrate the nature of these factors has been inconclusive. Subsequent studies have not substantiated in all cases the position taken by Gortner and Harris (*Bull. Torr. Bot. Club*, 1913, XL., 27). Studies on the osmotic concentration and the electrical conductivity of the fluids of the proliferous mass and of the wall have been given by Harris, Gortner and Lawrence, *Biochem. Bull.*, 1915, iv., 52.

If the formation of the basal prolification be due to the presence of special formative substances, one might occasionally expect to find the formation of carpellary tissue from other primordia. The only primordia normally developed subsequent to the carpels themselves are the ovules, which are borne on the carpellary margins.

To test this point, and to secure materials for other investigations, a series of dissections was begun in 1908. Those which were made from 1908 to 1915 are summarized in the accompanying table.

The results show that in the series of 568,098 dissections which have been made of fruits grown under a rather wide variety of conditions, basal prolification occurred 18,921 times, or in 3.330 per cent. of the fruits. Placental prolification occurred only 224 times or in .039 per cent. of the cases. Basal and placental prolification occurred in 18 of the 568,098 fruits.

While the occurrence of basal prolification presents a number of interesting morphological problems it does not seem to have the physiological significance of placental prolification. In the first case we have merely the continuation of activity of an axis which normally ceases with the laying down of the whorl of carpels forming the normal fruit. In the second case we have an entire transformation of a primordium. The primordium which should develop into an ovule forms instead a carpel, *i. e.*, one of the units of which the normal ovary is built up.

I am inclined to consider that this result is due to the local influence of special formative materials.

Experiment.	Without Prolification.	Basal Prolification.	Placental Prolification.	Basal and Placental Prolification.	Total Placental Prolification.	Total Fruits.	Percentage Placental Prolification.
1908	20,104	446	..	..	..	20,550	...
1909	116,821	4,622	30	..	30	121,473	.024
1911	30,105	873	9	..	9	30,987	.029
1912	10,487	441	1	..	1	10,929	.009
1913	123,216	4,458	17	3	20	127,694	.015
1914	180,516	7,143	144	6	150	187,809	.079
1915	67,686	938	23	9	32	68,656	.046
Total..	548,935	18,921	224	18	242	568,098	.042

2 (1584)

Formulæ for the determination of the correlations of size and of growth increments in the developing organism.

By J. Arthur Harris.

[*From the Station for Experimental Evolution, Cold Spring Harbor, L. I.*]

In the analysis of the growth of the higher organism it is essential to obtain definite measures of the interrelationship between certain measured magnitudes. Those which require consideration are the following:

(1) The correlations between the actual size of the organism at the various stages[1] of growth. (2) The correlations between growth increments of the organism during the several growth periods. (3) The correlations between the size of the organism at any stage and any or all subsequent growth increments.

The labor of determining these correlations by ordinary methods is excessive. If the first set of correlations (1) be determined by taking all moments about o as origin,[2] we may solve problems (2)–(3) as follows.

Problem 2.—To determine the correlations between growth increments from the moments and product moments of size at the several growth stages.

Let w, x, y, z be the dimensions of the organism at growth stages p, q, r, s. The growth increment during the intervals $q-p$, $r-q$, $s-r$ will then be $i_{pq} = x-w$, $i_{qr} = y-x$, $i_{rs} = z-y$.

---

[1] Growth stage denotes any given moment of time at which series of organisms of the same age are measured. During development it is, therefore, synonymous with age. The absolute size of the organism or any of its parts at a given growth stage is the only character of the organism available for consideration.

Growth period denotes the period of time elapsing between the sth and the $s + n$th growth stage, where s is any growth stage. Growth increment denotes the increase in size during any such period.

[2] Harris, J. Arthur, *Amer. Nat.*, 1910, xliv, 693–699.

The moments $\Sigma(x)$, $\Sigma(x^2)$, $\Sigma(y)$, $\Sigma(y^2)$, $\cdots$, and the product moments $\Sigma(wx)$, $\Sigma(wy)$, $\cdots$, $\Sigma(yz)$ are available for the correlations between size, which are required on their own account (Problem 1).

The constants for growth increments are given by well-known formulæ

$$\bar{i}_{pq} = \bar{x} \text{-} \bar{w}, \text{ etc.,}$$

$$\sigma^2_{i_{pq}} = [(\Sigma(w^2) + \Sigma(x^2) - 2\Sigma(wx))]/N - \bar{i}^2_{pq},$$

and similarly for $\sigma_{i_{qr}}$, $\sigma_{i_{rs}} \cdots$.

The product moment for any two growth increments, say i_{pq} and i_{rs}, is

$$\Sigma(i_{pq}i_{rs}) = \Sigma(wy) - \Sigma(wz) - \Sigma(xy) + \Sigma(xz).$$

In the special case in which three *consecutive* stages, say w, x, y, are involved we write

$$\Sigma(i_{pq}i_{qr}) = \Sigma(wx) - \Sigma(wy) + \Sigma(xy) - \Sigma(x^2).$$

Problem 3.—To determine the correlation between the size of the organism at any stage and any growth increment.

The notation of problem (2) may be used. The physical constants for the growth stages and growth increments have been given. The product moments are

$$\Sigma(wi_{pq}) = \Sigma(wx) - \Sigma(w^2), \Sigma(xi_{qr}) = \Sigma(xy) - \Sigma(x^2),$$

$$\cdots, \quad \Sigma(wi_{qr}) = \Sigma(wy) - \Sigma(wx), \quad \Sigma(wi_{rs}) = \Sigma(wz) - \Sigma(wy), \quad \text{etc.}$$

Illustrations of applicability will be given elsewhere.

INTER–PERIODIC CORRELATION IN THE ANALYSIS OF GROWTH.

J. ARTHUR HARRIS AND H. S. REED,

STATION FOR EXPERIMENTAL EVOLUTION, COLD SPRING HARBOR, LONG ISLAND.

I. INTRODUCTORY.

In the literature of growth, mathematical equations to describe changes in the actual size of the organism, or changes in the growth rate, are finding continuously widening applications. One has merely to refer to the papers by Robertson, Miyake, Moeser, Ostwald, Reed and Holland (1919), and Reed (1920)[1] for illustrations.

The criticism usually directed against such work is that in the higher organism, growth is a highly complex process, and that in consequence it cannot be represented mathematically. It is because of the very fact that growth is a complex process that mathematical analysis of the experimental data is necessary. Corollary to this must be the recognition of the fact that since growth is not a simple process, no one mathematical formula will be adequate for full description[2] and no one method adequate for complete analysis.

Our purpose in the present note is to illustrate on a series of data collected by one of us (1919) the application of inter-periodic correlation coefficients to certain phases of the problem of growth.

Before passing to the analysis, which is the special purpose of this paper, definition of the terms which will be used and a note

---

[1] Citations of literature may be traced from Reed's paper.

[2] Those who consider the possible adequacy of a single equation take the ground that if it be possible to represent the growth of an organism by a simple equation, it may be by virtue of the fact that during growth the various (often conflicting external) factors which affect the living substance are integrated by the organism.

on the nature of the data on which the statistical methods are illustrated are in order.

By growth stage we mean any given moment of time at which a series of organisms are measured. It is, therefore, synonymous with age during the growth period. The absolute size of the organism or of one or more of its parts at a given growth stage is the only character of the organism available for consideration.

By growth period we understand the period of time elapsing between the sth and the $s + n$th growth stage.

The increase in size during any such period we shall designate as a growth increment.

By relative growth increment, i_{rs}, we understand the ratio of the growth increment, i, to the absolute size of the individual at stage, r, where r and s are any two successive stages.

Turning now to the question of the original data as given in Table I. of Reed's (1919) publication we note from a study of the physical constants for absolute size in Tables I. and II. that there is an increase in the mean height of the plants up to the 77th day.

TABLE I.

STATISTICAL CONSTANTS FOR SIZE AT VARIOUS GROWTH STAGES.

Growth Stage.	Mean.	Standard Deviation.	Coefficient of Variation.
7	17.931	1.617	9.0
14	36.328	4.786	13.2
21	67.845	8.932	13.2
28	97.672	14.673	15.0
35	130.724	19.174	14.7
42	168.707	24.801	14.7
49	205.397	32.760	16.0
56	229.672	37.842	16.5
63	247.345	42.574	17.2
70	251.776	43.433	17.3
77	253.810	43.767	17.2

The increase from the 63d to the 70th and from the 70th to the 77th day is relatively slight, being only 4.43 cm. or 1.79 per cent. of the height for the 63d day in the first case and only 2.03 cm. or 0.81 per cent. of the value for the 70th day in the second case. The difference between the 84th day and the 77th day is negligible. In view of the fact that there is no appreciable growth in

the sense in which the term is used here between the 77th and the 84th day, this period will be left entirely out of account in the calculation of the correlations for the following discussions.

Furthermore by considering the constants for growth increments as shown in Table II., we note that the coefficients of varia-

TABLE II.

STATISTICAL CONSTANTS FOR GROWTH INCREMENTS FOR VARIOUS GROWTH PERIODS.

Growth Period.	Mean Increment.	Standard Deviation.	Coefficient of Variation.
7 to 14.	18.397	3.764	20.5
14 to 21.	31.517	5.164	16.4
21 to 28.	29.827	7.907	26.5
28 to 35.	33.052	7.505	22.7
35 to 42.	37.983	11.578	30.5
42 to 49.	36.690	14.266	38.9
49 to 56.	24.276	16.540	68.1
56 to 63.	17.672	13.803	78.1
63 to 70.	4.431	4.713	106.4
70 to 77.	2.034	5.096	250.5

tion for growth increments from the 63d to the 77th day are abnormally great. This may be in part due to biological causes, but it is doubtless due to a considerable extent to the relatively large error of measurement when the increment is very small in comparison with the size of the organism. If this be true, we should expect the correlations for actual size for the 63d to the 84th day to be about the same as those for the immediately preceding growth stages, but the correlations for growth increments may be expected to be of little value.

The problems which may be considered will be presented and discussed seriatim.

II. ANALYSIS OF DATA.

PROBLEM I. *The correlation between the absolute size of the organisms at its several periods of development.*

When examined at an early stage of development, organisms are found to differ among themselves in size. The same is found to be the case when the same series is measured at a later growth stage or at maturity.

In the biological analysis of the phenomenon of growth a prob-

lem of great importance is that of the causes which bring about the differences in size observable at any stage of development, or after growth has entirely ceased. Are individuals which are found to be small at maturity those which were small initially and have remained so from the beginning, or may the growth rate of an individual change during the course of its development to such an extent that it may vary its position in the series under investigation from time to time? That the latter is to some extent the case we know from general observations on human children. The problem to be solved is that of the quantitative magnitude of the relationship between the size of the individual at different stages of development.

The nature of the biological problems to be investigated has been stated in earlier work, and an attempt has been made to solve them by grouping plants according to quintile (Pearl and Surface, 1915) or quartile (Reed, 1919) position in the culture to which they belong and ascertaining the quartile or quintile in which they fall at different stages of growth.

This method has the disadvantage that all the individuals, whatever their size, are lumped together in four or five groups. In this method of treatment, small differences between two individuals are, therefore, given as much significance as large ones, providing they are large enough to throw the two individuals into different quartiles or quintiles.

An alternative method, which will completely obviate this difficulty, is to determine the correlation between the sizes of the individual at different periods of growth. The possible correlations between the absolute size of the individuals in the 11 different stages of growth of the *Helianthus* plants are shown in Table III.

The coefficients in this table can be best understood by first examining those for the relationships between the sizes of the plants near the period of maturity, and then passing to the relationships between the sizes of the plants at earlier stages.

Considering first of all the coefficients in the lower right-hand corner of the table, we note that all the coefficients are very high, denoting practically perfect correlation. This is the relationship which would be expected for a period when the organism has

TABLE III.

CORRELATION BETWEEN THE ACTUAL HEIGHT OF THE PLANTS AT THE VARIOUS GROWTH STAGES.

Stage.	7	14	21	28	35	42	49	56	63	70	77
7		+.733±.041 17.9	+.558±.061 9.14	+.468±.069 6.78	+.347±.078 4.46	+.193±.085 2.26	+.130±.087 1.50	+.093±.088 1.05	+.069±.088 0.78	+.065±.088 0.74	+.053±.088 0.60
14	+.733±.041 17.9		+.889±.019 48.0	+.695±.046 15.2	+.532±.064 8.38	+.343±.078 4.39	+.220±.084 2.60	+.151±.087 1.74	+.150±.087 1.73	+.140±.087 1.62	+.123±.087 1.41
21	+.558±.061 9.14	+.889±.019 48.0		+.887±.019 47.1	+.739±.040 18.4	+.552±.062 8.97	+.390±.075 5.20	+.320±.080 4.02	+.311±.080 3.90	+.297±.081 3.67	+.282±.082 3.46
28	+.468±.069 6.78	+.695±.046 15.2	+.887±.019 47.1		+.936±.011 85.2	+.752±.038 19.6	+.534±.063 8.44	+.409±.074 5.55	+.356±.077 4.61	+.329±.079 4.17	+.318±.080 3.99
35	+.347±.078 4.46	+.532±.064 8.38	+.739±.040 18.4	+.936±.011 85.2		+.892±.018 49.5	+.674±.048 13.9	+.488±.067 7.24	+.394±.075 5.26	+.350±.078 4.50	+.333±.079 4.23
42	+.193±.085 2.26	+.343±.078 4.39	+.552±.062 8.97	+.752±.038 19.6	+.892±.018 49.5		+.914±.015 62.5	+.732±.041 17.8	+.629±.053 11.8	+.580±.059 9.86	+.558±.061 9.16
49	+.130±.087 1.50	+.220±.084 2.60	+.390±.075 5.20	+.534±.063 8.44	+.674±.048 13.9	+.914±.015 62.5		+.900±.017 53.5	+.819±.029 28.1	+.775±.035 21.9	+.758±.038 20.1
56	+.093±.088 1.05	+.151±.087 1.74	+.320±.080 4.02	+.409±.074 5.55	+.488±.067 7.24	+.732±.041 17.8	+.900±.017 53.5		+.948±.009 105	+.926±.013 73.3	+.911±.015 60.2
63	+.069±.088 0.78	+.150±.087 1.73	+.311±.080 3.90	+.356±.077 4.61	+.394±.075 5.26	+.629±.053 11.8	+.819±.029 28.1	+.948±.009 105		+.994±.001 970	+.986±.002 411
70	+.065±.088 0.74	+.140±.087 1.62	+.297±.081 3.67	+.329±.079 4.17	+.350±.078 4.50	+.580±.059 9.86	+.775±.035 21.9	+.926±.013 73.3	+.994±.001 970		+.993±.001 828
77	+.053±.088 0.60	+.123±.087 1.41	+.282±.082 3.46	+.318±.080 3.99	+.333±.079 4.23	+.558±.061 9.16	+.758±.038 20.1	+.911±.015 60.2	+.986±.002 411	+.993±.001 828	

practically attained its adult size and in which there is relatively little change from one week to another.

As we follow the correlations between the later periods and preceding periods back, we note that there is a regular decrease in the values of the correlation coefficients. This may be best shown by summarizing the results graphically in diagram 1.

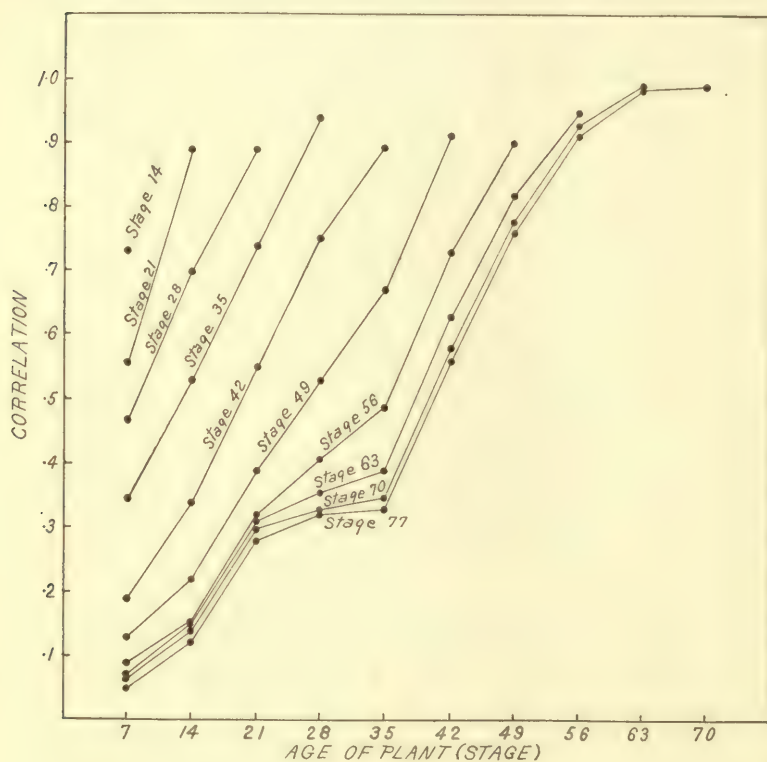In the graph the correlation of the size of the organism at each

DIAGRAM 1.

growth stage with its size at every antecedent growth stage (shown at the bottom of the diagram) is shown on the scale of correlation at the left by points marking the magnitude of the correlations for each of the growth stages. The pitch of the lines connecting the points for the 14th to the 77th growth stage shows the rapid decrease in the magnitude of the correlations as the stages become more widely separated in time.

The same type of diagram may be used to show the relation-
ship between the size at early and at later growth stages. Dia-
gram 2 shows the distribution of the magnitudes of the correla-
tions for sizes of the individuals at the 7th to the 70th day (stage)
and the size at subsequent growth stages.

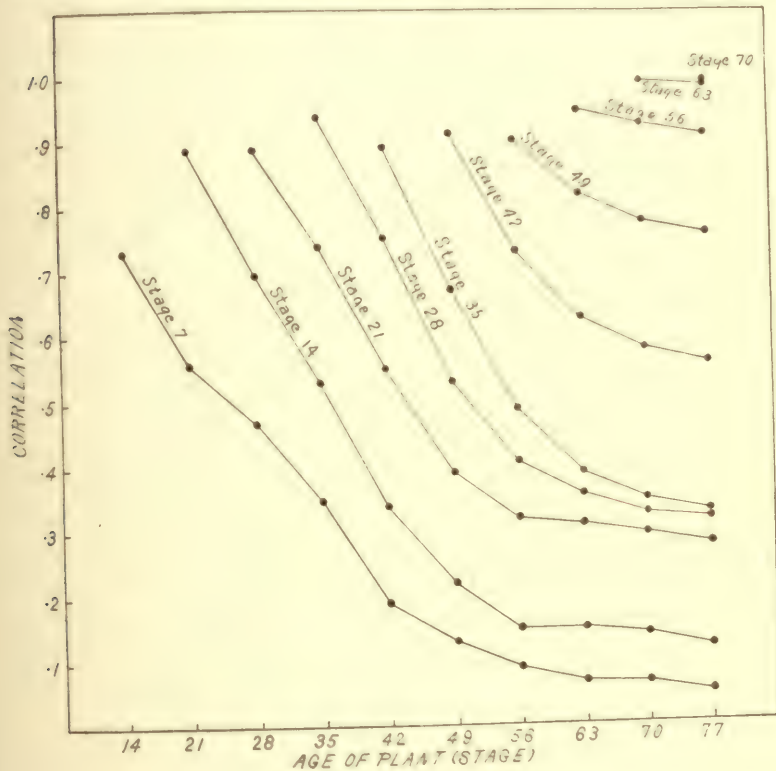From these lines it is clear that the correlations between size

DIAGRAM 2.

at antecedent and subsequent periods decrease as the periods be-
come more widely separated in time. This is true without ex-
ception for every period which furnishes evidence upon the
question.

The coefficients are, however, positive in sign throughout, thus
suggesting (though in some cases not proving) that throughout
its growth period the size of the plant bears some relation to its
size when first measured. This result is in agreement with the

findings of Webber (1920) in regard to the growth of *Citrus* stock.

PROBLEM 2. *The correlation between the growth increments of the organism during the several growth periods.*

Our second problem is to determine whether there is a correlation in growth increments as well as in actual size of the organism. We shall thus answer the question whether the organism which grows more rapidly than the average during one growth period will grow more rapidly than the average in other growth periods and whether the organism which lags behind the average in its rate of growth during one growth period will also lag behind during other growth periods.

Little has heretofore been done towards the statistical treatment of growth increments. This is probably in part due to the arithmetical difficulties of computing the constants for increments, but if the moments and product moments be taken about zero as origin in computing the coefficients required under Problem 1 above, the calculations for growth increments are easily made by the use of formulæ given elsewhere (Harris, 1920).

The symmetrical table showing the relationship between the actual growth increments for all of the combinations of growth periods appears as Table IV. This table shows positive and statistically significant correlation coefficients for closely associated periods throughout the season up to and including the period for the 63d to the 70th day. The coefficients for the period from the 70th to the 77th day cannot in general be considered statistically significant in comparison with their probable errors.

Examining these results in a little greater detail, we note that the nine coefficients showing the relationship between the growth increments of successive weeks (the constants bordering the diagonal cell of the symmetrical table of constants) are all positive in sign and with the exception of the last (showing the relationship between the growth of the period from the 63d to 70th and that between the 70th to 77th day) all are statistically significant. The eight coefficients measuring the correlations between the growth increments of weekly periods which are separated by one week are also without exception positive, but are lower in magnitude and less certainly statistically significant. For periods more

TABLE IV.

CORRELATIONS BETWEEN THE GROWTH INCREMENTS DURING THE SEVERAL GROWTH PERIODS.

Growth Period.	Growth Period.									
	7 to 14.	14 to 21.	21 to 28.	28 to 35.	35 to 42.	42 to 49.	49 to 56.	56 to 63.	63 to 70.	70 to 77.
7 to 14...		+.655 ±.051 12.9	+.259 ±.083 3.14	+.013 ±.089 0.14	−.115 ±.087 1.32	−.102 ±.088 1.17	−.095 ±.088 1.08	+.082 ±.088 0.93	−.070 ±.088 0.79	−.136 ±.087 1.57
14 to 21...	+.655 ±.051 12.9		+.630 ±.053 11.8	+.264 ±.082 3.21	+.064 ±.088 0.73	−.024 ±.089 0.27	+.011 ±.089 0.13	+.100 ±.088 1.14	−.082 ±.088 0.93	−.047 ±.088 0.53
21 to 28...	+.259 ±.083 3.14	+.630 ±.053 11.8		+.636 ±.053 12.1	+.161 ±.086 1.86	−.079 ±.088 0.90	−.179 ±.086 2.08	−.140 ±.087 1.61	−.249 ±.083 2.99	−.023 ±.089 0.26
28 to 35...	+.013 ±.089 0.14	+.264 ±.082 3.21	+.636 ±.053 12.1		+.532 ±.064 8.37	+.147 ±.087 1.70	−.316 ±.080 3.96	−.273 ±.082 3.33	−.483 ±.068 7.12	−.169 ±.086 1.97
35 to 42...	−.115 ±.087 1.32	+.064 ±.088 0.73	+.161 ±.086 1.86	+.532 ±.064 8.37		+.778 ±.035 22.3	+.070 ±.088 0.79	+.067 ±.088 0.76	−.189 ±.085 2.21	−.106 ±.088 1.21
42 to 49...	−.102 ±.088 1.17	−.024 ±.089 0.27	−.079 ±.088 0.90	+.147 ±.087 1.70	+.778 ±.035 22.3		+.416 ±.073 5.67	+.250 ±.083 3.01	−.005 ±.089 0.06	+.038 ±.088 0.43
49 to 56...	−.095 ±.088 1.08	+.011 ±.089 0.13	−.179 ±.086 2.08	−.316 ±.080 3.96	+.070 ±.088 0.79	+.416 ±.073 5.67		+.299 ±.081 3.70	+.453 ±.070 6.44	+.022 ±.089 0.25
56 to 63...	+.082 ±.088 0.93	+.100 ±.088 1.14	−.140 ±.087 1.61	−.273 ±.082 3.33	+.067 ±.088 0.76	+.250 ±.083 3.01	+.299 ±.081 3.70		+.476 ±.069 6.91	+.188 ±.085 2.20
63 to 70...	−.070 ±.088 0.79	−.082 ±.088 0.93	−.249 ±.083 2.99	−.483 ±.068 7.12	−.189 ±.085 2.21	−.005 ±.089 0.06	+.453 ±.070 6.44	+.476 ±.069 6.91		+.086 ±.088 0.97
70 to 77...	−.136 ±.087 1.57	−.047 ±.088 0.53	−.023 ±.089 0.26	−.169 ±.086 1.97	−.106 ±.088 1.21	+.038 ±.088 0.43	+.022 ±.089 0.25	+.188 ±.085 2.20	+.086 ±.088 0.97	

widely separated in time the correlations are in part positive and in part negative in sign.

Thus from the results as a whole it appears that the increments of successive periods are generally positive and fairly highly correlated when the periods show actual growth increments. Thus the zone of coefficients lying along the diagonal cell are positive and generally fairly high. When the periods are separated by any considerable length of time the coefficients are generally insignificant in magnitude and may, as a matter of fact, be either positive or negative in sign.

The relationship may be brought out by determining the averages of the correlation coefficients, with regard to sign, for the increments of periods separated by various lengths of time. The results are as follows.

Period of Separation (Weeks).	Number of Correlations Averaged.	Average Correlation.
0	9	+ .5009
1	8	+ .2240
2	7	— .0334
3	6	— .1236
4	5	— .1640
5	4	— .1033
6	3	— .0077
7	2	— .0585
8	1	— .1360

If we disregard the cases in which there are less than five coefficients to be averaged, we note a steady decrease in the magnitude of the correlation coefficient. Periods of growth which are successive or separated by only one week have positively correlated growth increments. Periods which are more widely separated show negative correlations of the increments.

The relationship between the coefficients in Table IV. may be clarified by diagram 3 which shows the relationship between four of the ten growth increments and each of the other ten increments. The increments selected as a "first variable" in the correlation are the first, fourth, seventh, and tenth. This has the advantage of representing the first and the last growth increment, and of leaving undrawn no more than two successive increments. The figures are aligned according to the ten increments representing the "second variable" of the correlation.
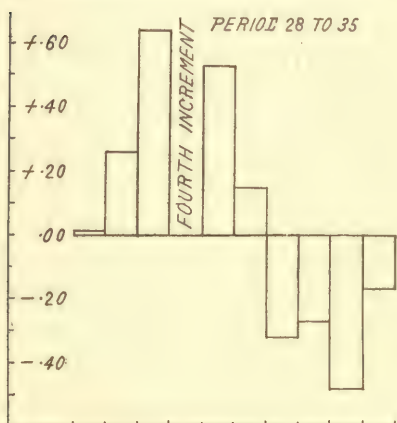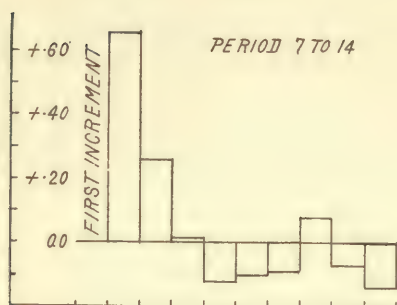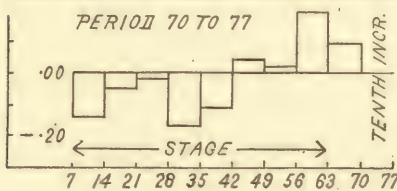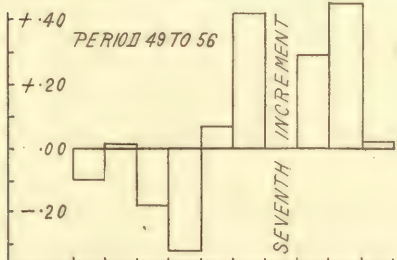
DIAGRAM 3.

The graphs for the first, fourth and seventh increment show clearly the shift in the position of the maximum positive correlation from the earlier to the later periods as the "first variable" is chosen from the later periods. The same is shown less clearly by the correlations for the tenth increment, but there the coefficients are very small, presumably because growth has practically ceased.

It is clear, therefore, that plants which are growing more rapidly during any period of development will grow more rapidly during a closely associated subsequent period of development but that there is little or no relationship, or even a negative relationship, between the rate of growth of the organisms studied at considerably separated periods of time.

Since the correlations for absolute growth increments are so small for all except successive periods of time, it seems unnecessary to deal at present with the relative growth increments, *i.e.*, with the growth increments expressed as a fraction of the size of the organisms at the beginning of the growth period.

PROBLEM 3. *The correlation between the absolute size of the organism at given stages of development and subsequent growth increments.*

In the higher plant organism rate of growth at any period must be supposed to depend to some extent upon plastic materials synthesized by the more nearly mature portions of the same individual. Thus one might expect to find a relationship between the actual size of the organism at any stage of growth and the rate at which the organism increases in size during a subsequent period.

We have determined the possible correlations between the absolute size of the organism at different periods and the growth increment of the organism during subsequent growth periods. The coefficients are presented in Table V. This shows positive correlation between the actual size of the organism at every stage of development from the 7th to the 70th day and the increase in the size of the organism during the following week. The magnitude of the correlation is of the order $r = 0.45$ to $r = 0.60$ for the 7th, 14th, 21st, and 28th day. For these growth stages the correlation between actual size and the subsequent growth incre-

TABLE V.

Correlations between Height at Each Growth Stage and the Increments at Subsequent Growth Periods.

Stage	Growth Period									
	7 to 14.	14 to 21.	21 to 28.	28 to 35.	35 to 42.	42 to 49.	49 to 56.	56 to 63.	63 to 70.	70 to 77.
7	+.503 ±.066 / 7.60	+.285 ±.081 / 3.50	+.439 ±.084 / 2.86	−.028 ±.089 / 0.32	−.162 ±.086 / 1.88	−.036 ±.088 / 0.41	−.046 ±.088 / 0.52	−.042 ±.088 / 0.47	−.021 ±.088 / 0.24	−.100 ±.088 / 1.14
14		+.611 ±.056 / 11.0	+.285 ±.081 / 3.50	+.000 ±.089 / 0.01	−.146 ±.087 / 1.68	−.093 ±.088 / 1.06	−.090 ±.088 / 1.03	+.050 ±.088 / 0.57	−.062 ±.088 / 0.70	−.141 ±.087 / 1.63
21			+.517 ±.065 / 7.96	+.153 ±.087 / 1.77	−.041 ±.088 / 0.46	−.063 ±.088 / 0.72	−.042 ±.088 / 0.48	+.084 ±.088 / 0.96	−.081 ±.088 / 0.92	−.103 ±.088 / 1.18
28				+.436 ±.072 / 6.08	+.062 ±.088 / 0.70	−.081 ±.088 / 0.92	−.122 ±.087 / 1.40	−.024 ±.089 / 0.27	−.183 ±.086 / 2.14	−.075 ±.088 / 0.85
35					+.255 ±.083 / 3.09	−.004 ±.089 / 0.05	−.217 ±.084 / 2.57	−.125 ±.087 / 1.43	−.329 ±.079 / 4.17	−.124 ±.087 / 1.42
42						+.360 ±.077 / 4.67	−.135 ±.087 / 1.56	−.065 ±.088 / 0.74	−.343 ±.078 / 4.39	−.145 ±.087 / 1.67
49							+.079 ±.088 / 0.89	+.059 ±.088 / 0.67	−.262 ±.083 / 3.17	−.093 ±.088 / 1.06
56								+.182 ±.086 / 2.12	−.029 ±.089 / 0.32	−.071 ±.088 / 0.81
63									+.129 ±.087 / 1.48	−.002 ±.089 / 0.03
70										+.007 ±.089 / 0.08

ment is clearly significant in comparison with its probable error. The coefficients are lower for the 35th and the 42d day, but are probably statistically trustworthy. Beyond this period there seems to be no relationship between the size of the organism and its growth rate in an immediately following period.

For the first two stages of growth measured, the 7th and the 14th day, there may be a significant correlation of the order $r = + .285$ between size and growth increments during the second following week.

The coefficient of correlation between size and the increment in the second week following is also positive for the 21st and 28th day, but neither of these values may be considered statistically significant in comparison with its probable error. Finally for the first stage (seventh day) there may be a significant correlation between absolute size and growth increments during the third week following ($r = + .239$ for increment for 21st to 28th day). Other than this the coefficients are for the most part statistically insignificant in comparison with their probable error.

Summarizing the preceding statements as a basis for further analysis, we note that for the first six growth stages (7th to 42d day) there is a significant positive correlation between the size of the organism at the given stage and the growth increment of the following week. For the first two growth stages (and possibly in the third where $r/E_r = 1.77$) there is a significant correlation between the size of the organism and the growth increment in the second subsequent week. Finally, for the first stage only, there is a significant positive correlation between size and growth increment in the third subsequent week.

Disregarding these 9 coefficients and the 4 positive but not significant correlations between the sizes at the several growth stages and the growth increments of the following week, we may note the following facts concerning the remaining 42 coefficients.

Of these 42 coefficients 36 are negative while only 6 are positive in sign. Of the 6 positive coefficients only that between actual size on the 21st day and growth increment between the 28th and 35th day (already considered above) is as large as its probable error. Of the 36 negative coefficients 18 are larger than

their probable error, and 5 of these are over twice as large as their probable error.

There is, therefore, clear evidence that the subsequent growth of the higher plant organism is measurably conditioned by its size. In general the larger individuals grow more rapidly in immediately subsequent periods, but somewhat more slowly than the average in more distant subsequent periods.

While a detailed discussion of the relation of these results to the theory that growth may be satisfactorily described by the curve of an autocatylic reaction falls quite outside the scope of this paper, it must be noted that negative correlations between actual size at a given stage and the growth increments of certain subsequent growth periods might be expected. As Reed and Holland (1919) have pointed out the plants attained about half their final height at about the thirty-fourth day. From this time on the increments were decreasing. Plants which had attained more than the average size at this period would, therefore, of necessity, make smaller average increase in size in later periods.

The number of individuals measured is not sufficiently large to carry the analysis farther.

III. Recapitulation.

The purpose of this paper has been to illustrate on the basis of a specific series of data the value of the inter-periodic correlation coefficient in the analysis of the phenomena of growth.

The analysis shows that in the case of a series of *Helianthus* plants the actual size of the individual at any stage of development is closely correlated with its size at other closely associated stages of development. The magnitude of the correlation rapidly diminishes as the growth stages become more widely separated in time. Thus the final size of the organism is but to a slight extent determined by its initial size.

The correlation between successive growth increments is positive in sign and statistically significant, with the general average of $\bar{r} = .501$. The correlation for increments of weekly periods separated by one week is on the average only about $\bar{r} = .225$. For periods more widely separated than this the correlation between growth increments is on the average negative in sign.

Thus plants which are growing more rapidly during one period of development will grow more rapidly during a closely associated period, but there is little or no relationship between the growth increments of more widely separated periods.

The growth increment of the organism is positively correlated with its size at an immediately preceding stage. In the early stages of growth, the growth increments of two or even three subsequent periods are positively correlated with the initial size of the organism.

LITERATURE CITED.

Harris, J. Arthur.

'20   Formulæ for the Calculation of the Correlations of Size and of Growth Increments in the Developing Organism. Proc. Soc. Exp. Biol. and Med. *In press.*

Pearl, R., and Surface, F. M.

'15   Growth and Variation in Maize. Zeitschr. Indukt. Abst. u. Vererbungslehre, **14**: 97–203.

Reed, H. S.

'19   Growth and variability in *Helianthus*. Amer. Jour. Bot., 252–271.

'19   The Growth Rate of an Annual Plant, *Helianthus*. Proc. Nat. Acid. Sci., **5**: 135–144.

Reed, H. S.

'20   The Nature of the Growth Rate. Jour. Gen. Physiol., **2**: 545–561.

Webber, H. J.

'20   Selection of Stocks in *Citrus* propagation. Bull. Univ. Calif. Coll. Agr., 317

NOTES ON THE OCCURRENCE OF GAMMERUS LIMNAEUS SMITH IN A SALINE HABITAT

THE capacities of various organisms for withstanding relatively wide ranges of environmental conditions has received considerable attention at the hands of physiologists and students of animal behavior, and is a problem which must ultimately be considered in greater detail by ecologists, students of geographic distribution and organic evolution.

The purpose of this note is merely to call attention to the occurrence of *Gammerus limnaeus* Smith,[1] normally a fresh water[2]

[1] We are indebted to Mr. Waldo L. Schmitt, associate curator of marine invertebrates in the U. S. National Museum, for the determination of the species. The specimens are in the National Museum.

[2] The key to the taxonomic and distributional literature is furnished by Weckel's paper on the fresh water Amphiopoda of North America (*Proc. U. S. Nat. Mus.*, 32: 42–44, 1907), and individual citations need not be given here. The species was first dredged in Lake Superior. It has been taken near Long's Peak, Colorado, at an elevation of 9,000 feet; from a cool spring, Fire Hole Basin; from Shoshone Falls, Idaho; Flathead Lake, Montana; and from the Yellowstone National Park. It is reported from Fort Wingate, N. M., and from the Wasatch Mountains and Salt Lake City, Utah. It is impossible to determine from the records whether all the localities were fresh water habitats, but that it is typically a fresh water form can admit of no possible doubt. It has been taken from the stomachs of trout from brooks near Marquette, Mich.

species, in a peculiar and rather saline habitat.[3]

In the summer of 1920 the writers visited the Ice Spring Craters lava field of the Sevier Desert in the ancient Lake Bonneville basin described in detail by Gilbert.[4] On climbing down into the old lava vent[5] of the Terrace crater we were surprised to find a small crustacean abundant in the small pool of clear water at the bottom. It was noted that a number of the animals were very slightly pigmented, apparently indicating that in the semi-darkness of the pool they were approaching cave conditions. In all instances, however, the eyes were fully pigmented. The presence of the *Gammerus* led to the assumption that the water was non-saline and we were preparing to replenish our water bag when taste showed it to be distinctly brackish.

A sample of the water was therefore taken in a clean Mason fruit jar from which it was afterwards transferred to citrate bottles for shipment to the laboratory. The water had a freezing point lowering of 0.410° C., indicating an osmotic concentration of 4.94 atmospheres and an electrical conductivity of .0138 reciprocal ohm. The hydrogen ion concentration of the water (determined electrometrically) was $C_H = 0.409 \times 10^{-7} = p^H 7.388$. Analysis showed the following composition.

[3] The genus *Gammerus* has species which occur in more or less saline coastal habitats and in non-saline inland waters.

[4] Gilbert, G. K., "Survey West of the 100th Meridian," Vol. 3, pp. 136–144; also "Lake Bonneville," Monographs U. S. Geol. Survey, I., pp. 320–325, 1890.

[5] The lava vent is a circular tube, at one side of the wide crater, about 12 feet in diameter inclined 10° or 15° from the vertical. It can be explored for about 25 feet when progress is stopped by water.

Grams per Liter

	Grams per Liter
Total solids (at 110°)..	8.5666
Total solids (at 210°)..	8.1467
Total solids (ignited)[6]..	7.6400
CO_3[7]	none
HCO_3[7]	0.2187

Mineral Analysis

	Grams per Liter	Per Cent. of Total Solids (Ignited)
SiO_2	0.0720	0.94
$Fe_2O_3Al_2O_3$	0.0030	0.04
Ca	0.3305	4.33
Mg	0.2560	3.35
Na	1.9750	25.85
K	0.3050	3.99
Cl	3.4120	44.66
SO_4	1.3260	17.36
CO_3[8]	0.1075	1.41
Total	7.7870	101.93

Hypothetical Inorganic Composition of the Solution

	Grams per Liter	Per Cent. of Total
Na_2SiO_3	0.1460	1.84
$Ca(HCO_3)_2$	0.2913	3.68
$CaSO_4$	0.8780	11.08
$MgSO_4$	0.8855	11.18
$MgCl_2$	0.3023	3.81
KCl	0.5875	7.42
NaCl[9]	4.8330	60.99
Total	7.9236	100.00

[6] There was apparently considerable organic matter in solution. This could easily be derived from bat guano which was observed on the lava ledges surrounding the pool.

[7] Carbonates and bicarbonates were determined by the titrametric method proposed by Scales (SCIENCE, N. S., 51, p. 214, 1920).

[8] Calculated from bicarbonate data according to the formula $2RHCO_3 + heat = R_2CO_3 + CO_2 + H_2O$.

[9] An average value based on NaCl contents of 4.8790 gr. calculated from residual Na and 4.7870 calculated from residual Cl. The difference of 0.092 gram per liter is within experimental error when one remembers that the above calculations are

The Terrace crater, and indeed all of the craters of the Ice Spring Craters group, is unquestionably post-Bonneville in origin. There is no trace of wave work on the outer slopes of the craters such as are so conspicuous on Pavant Butte to the north, and neither lacustrine sediments nor evidences of subaqueous erosion appear on the surface of the evidently recent lava fields as they do on the Fumarole Butte lava field to the northwest.

The depth of the vent of the Terrace crater is 260 feet below its general rim and 220 feet below the sill of the last outflow. The problem of the original introduction of *Gammerus* into the small pool of water occupying the bottom of this crater is that of the transportation of small crustacean species or their eggs in general. The point of physiological interest is the occurrence of this species, hitherto reported from non-saline waters, in water of this concentration.

Ross Aiken Gortner,
Division of Agricultural Biochemistry, University of Minnesota,

J. Arthur Harris,
Station for Experimental Evolution, Carnegie Institution of Washington

purely empirical and also when one considers that in some instances the actual analytical values, and consequently accompanying experimental errors, were multiplied by 50 to bring the calculation to a liter basis.

SECONDARY PARASITISM IN PHORADENDRON

BROWN'S[1] illustration of *Phoradendron californicum* parasitic on *P. flavescens*[2] has a twofold interest. First, it records a case of secondary parasitism which seems to be very rare indeed. It has never, so far as I am aware, been noted by workers at the Desert Botanical Laboratory, a number of whom have been especially interested in parasitism. For the most part *P. macrophyllum* and *P. californicum* occur on quite different hosts.[3] Second, the case is of interest physiologically, as BROWN suggests, in its relation to osmotic and other physical phenomena. HARRIS and LAWRENCE, in their study of the sap properties of Jamaican montane rain forest Loranthaceae,[4] find that in these forms the sap extracted from the green stems of the leafless species shows lower osmotic concentration than that from the foliar tissues of the leafy forms. Thus in working with 7 species of Loranthaceae they found average values of the freezing point lowering of $1.153°$, $1.176°$, and $1.177°$ in the leafless species as compared with $1.305°$, $1.347°$, $1.400°$, and $1.650°$ in

[1] BROWN, J. G., Mistletoe vs. mistletoe. BOT. GAZ. **65**:193. *fig. 1.* 1918.

[2] This is presumably *P. macrophyllum* Cockerell, the *P. flavescens macrophyllum* of ENGLEMANN and of some subsequent workers, or one of its varieties. The host here, as Professor BROWN has kindly written me, was a *Fraxinus*.

[3] TRELEASE (The genus *Phoradendron*, p. 14, Urbana. 1916) notes that *P. californicum*, while occurring exclusively on angiosperms, belongs to a group, the "Pauciflorae," which with this and one other exception is limited to coniferous hosts.

[4] HARRIS, J. ARTHUR, and LAWRENCE, J. V., On the osmotic pressure of the tissue fluids of Jamaican Loranthaceae parasitic on various hosts. Amer. Jour. Bot. **3**:438–455. 1916.

the leaves of the leafy forms. If the same is true of desert Loranthaceae, the relationship between leafless and leafy parasite observed by BROWN is just the reverse of what might be expected if successful parasitism were dependent upon higher osmotic concentration in the tissue fluids of the parasite.

As pointed out elsewhere, however, the technical difficulties in the comparison of the tissue fluids of the stems and leaves by the methods as yet available for field work are rather great. In the leafless forms there is danger of including a considerable amount of fluids from woody conducting tissue not at all comparable with that of the green tissue which may be taken to be physiologically homologous with the leaf tissue of the leaves of the tree or of the leafy Loranthaceae. Furthermore, such work as has been done on the rather difficult problem of the physico-chemical properties of the tissue fluids of desert Loranthaceae[5] is insufficient to show that the osmotic concentration is lower in the leafless desert forms. Furthermore, the concentration of the sap of desert forms seems to vary rather widely, and even if the average concentration of the fluids of *P. californicum* were lower than that of *P. macrophyllum*, it is quite possible that the individual secondary parasite, *P. californicum*, had a higher concentration than its individual *P. macrophyllum* host.[6]

So far as I am aware, the only direct determination of osmotic concentration in primary and secondary parasitism in the Loranthaceae is that by HARRIS and LAWRENCE (*loc. cit.*) on the Jamaican broad-leaved *Phthirusa parvifolia* parasitic upon the leafless *Dendrophthora gracilis*, which is in turn parasitic upon a tree, *Cyrilla racemiflora*. The sap properties stand in the following relationship: *Cyrilla racemiflora*, $\Delta = 1.18$, $P = 14.2$; *Dendrophthora gracilis* (on *Cyrilla racemiflora*), $\Delta = 1.26$, $P = 15.2$; *Phthirusa parvifolia* (on *Dendrophthora gracilis*), $\Delta = 1.49$, $P = 17.9$. Osmotic concentration increases from the host to the primary parasite and from the primary parasite to the secondary one. Note also that the observed secondary parasitism is the *leafy P. parvifolia* with an average depression of $1.347°$ upon the *leafless D. gracilis* with an average depression of $1.176°$.—J. ARTHUR HARRIS, *Cold Spring Harbor, N.Y.*

[5] HARRIS, J. ARTHUR, On the osmotic concentration of the tissue fluids of desert Loranthaceae. Mem. Torr. Bot. Club **17**:307–315. 1918.

[6] I have individual determinations on *P. californicum* which indicate higher concentration than some found in *P. macrophyllum*. The great difficulty of comparing the sap properties of the two forms lies in the fact that, in the neighborhood of Tucson at least, they occur in the main on different hosts and for the most part in slightly different local habitats.

THE INTERRELATIONSHIP OF THE NUMBER OF STAMENS AND PISTILS IN THE FLOWERS OF *FICARIA*.

J. ARTHUR HARRIS.

THE INTERRELATIONSHIP OF THE NUMBER OF STAMENS AND PISTILS IN THE FLOWERS OF *FICARIA*.

J. ARTHUR HARRIS.

I. INTRODUCTORY REMARKS.

A survey of the rapidly increasing literature must convince anyone that the problem of the factors which determine the sex of the organism is one of such complexity that it cannot be solved on the basis of any one kind of material or by any one method of research.

In the flowering plants the same individual may produce both eggs and sperm. The relative numbers of egg and sperm producing organs may vary from individual to individual, or from flower to flower within the individual.

It is reasonable to assume that definite genetic, morphogenetic or physiological factors underlie these variations. Any successful attempt to determine these factors and to measure their influence is just as truly a contribution to the wide problem of the physiology of sex as the more conventional breeding experiments and studies on the morphology of the germ cells.

The purpose of this paper is to point out certain hitherto unrecognized relationships between the number of sporophylls in the flower of the ranunculaceous genus *Ficaria*.

Heretofore those who have investigated the problem of the relationship between the number of stamens and pistils in the flower have been content to merely determine the correlation between the number of the two kinds of spore-bearing organs. Positive correlations of this kind should arise as the resultant of any sets of environmental factors which tend to increase both the number of stamens and the number of pistils in certain of the plants or individual flowers and to limit the number of both of these organs in others. Morphogenetically and physiologically it seems of far greater importance to inquire whether the relative

7

proportion of the two types of spore bearing organs is correlated with the total number of sporophylls, which in lieu of any better character may serve as a measure of the total influence of intrinsic and extrinsic factors influencing degree of development.

Several years ago Professor Pearson and I (Harris, '09) showed that problems of this kind can be approached by determining the correlation between the total organs laid down and the deviation of the number of a particular kind from the probable number on the assumption that the proportion of the particular kind is independent of the total number.

The statistical method may of course be applied to experimental data or to series of determinations made on organisms developing under natural conditions. As yet experimental series are not available.

In a former paper ('16) I showed that in the inflorescences of both *Arisarum vulgare* and *A. proboscidium* there is a significant negative correlation between the total number of flowers and the deviation of the number of staminate flowers from their probable number on the theory of proportional distribution. Thus the male flowers while *absolutely* more numerous in the inflorescences with larger total numbers of flowers are *relatively* less numerous than in the inflorescences with smaller total numbers of flowers. Or, conversely, the larger inflorescences tend to produce *a larger proportion* of pistillate flowers.

In this paper the same analytical methods will be applied to the problem of the relationship of the number of stamens and the number of pistils to the total number of stamens and pistils produced by the flower.

II. MATERIALS.

The materials upon which the coefficients discussed in this paper are based have been tabled and the chief biometric constants deduced by competent statisticians. The special methods upon which the conclusions of this paper are based were not, however, available at the time their calculations were made. The results are, therefore, quite new.

The series employed are the following:

1–2. A series of 283 countings of number of stamens and

pistils of *Ficaria verna* from Trogen and another series of 80 countings from Gais, published by Ludwig ('01). Statistical constants for both of these series have been deduced and published by Dr. Alice Lee ('02).

3–4. A series of 268 early and 373 late flowers of *Ficaria ranunculoides* collected by MacLeod ('99) and discussed by W. F. R. Weldon ('01).

5–8. Four series of *Ficaria ranunculoides* collected by Galton, Weldon, Pearson ('03) and others in Italy, Guernsey and England.

III.   PRESENTATION OF DATA.

The means and variabilities of number of stamens and pistils per flower have been given in the papers cited. The only point which requires discussion in this place is the relative variability of the number of the two types of sporophylls. This is shown in Table I.

TABLE I.

RELATIVE VARIABILITIES IN NUMBER OF STAMENS AND NUMBER OF PISTILS IN *Ficaria*.

Series.	Number of Flowers.	Coefficient of Variation for Pistils.	Coefficient of Variation for Stamens.	Differences
Switzerland—				
Trogen, I	283	23.32	18.68	4.64
Gais, II	80	23.73	12.18	11.55
Belgium—				
Early, III	268	23.32	14.07	9.25
Late, IV	373	27.89	18.46	9.43
Italy, V	624	22.35	14.12	8.23
Guernsey, VI	520	26.54	17.16	9.38
England—				
Dorset, VII	505	26.38	16.84	9.54
Surrey, VIII	500	27.19	17.32	9.87

The number of pistils is consistently more variable than the number of stamens.

Other workers have shown that there is a correlation of medium intensity between the number of stamens and the number of pistils per flower. Their constants, all of which have been rechecked in the course of this work, are shown in Table II. I have also added the linear regression equations showing the rate of increase in mean number of pistils associated with an increase in the number of stamens and the rate of increase in

TABLE II.

CORRELATION BETWEEN NUMBER OF STAMENS AND PISTILS IN *Ficaria* AND RE-
GRESSION EQUATIONS FOR STAMENS AND PISTILS.

Series.	Number of Flowers.	Correlation Stamens and Pistils.	Ratio of Correlation to Probable Error r/E_r.	Regression Line.
Switzerland—				
Trogen..........	283	.530 ± .029	18.27	$S = 11.708 + .645\,P$
				$P = 4.461 + .429\,S$
Gais.............	80	.388 ± .064	6.06	$S = 19.075 + .262\,P$
				$P = 4.403 + .575\,S$
Belgium—				
Early............	268	.507 ± .031	16.35	$S = 18.197 + .489\,P$
				$P = 3.427 + .525\,S$
Late.............	373	.749 ± .015	49.93	$S = 9.006 + .729\,P$
				$P = -1.593 + .769\,S$
Italy.............	624	.439 ± .022	19.95	$S = 19.313 + .418\,P$
				$P = 5.409 + .460\,S$
Guernsey..........	520	.534 ± .021	25.43	$S = 18.302 + .404\,P$
				$P = 4.160 + .707\,S$
England—				
Dorset...........	505	.669 ± .017	39.35	$S = 20.369 + .535\,P$
				$P = -1.333 + .835\,S$
Surrey...........	500	.660 ± .017	38.82	$S = 19.091 + .588\,P$
				$P = -.860 + .741\,S$

mean number of stamens associated with an increase in number of pistils per flower.

The regression of the number of stamens on the number of pistils and of the number of pistils on the number of stamens is shown for three of the larger series in Figs. 1–3.

The empirical means for the Italian series, Diagram 1, do not conform very satisfactorily to the lines given by the equations. Better agreements between actual and theoretical means could hardly be found (in series of data no larger than these) than in the Guernsey and Surrey series represented in Figs. 2 and 3.

The main purpose of the present paper is to present the results of the determination of the relationship between the total number of sporophylls and the number of stamens and pistils.

The correlations between the total number of sporophylls and the number of stamens and pistils are shown in Table III.

As is to be expected the correlations between total sporophylls and number of stamens and pistils are high.

The constants showing the relationship between the total number of sporophylls and the deviation of the number of
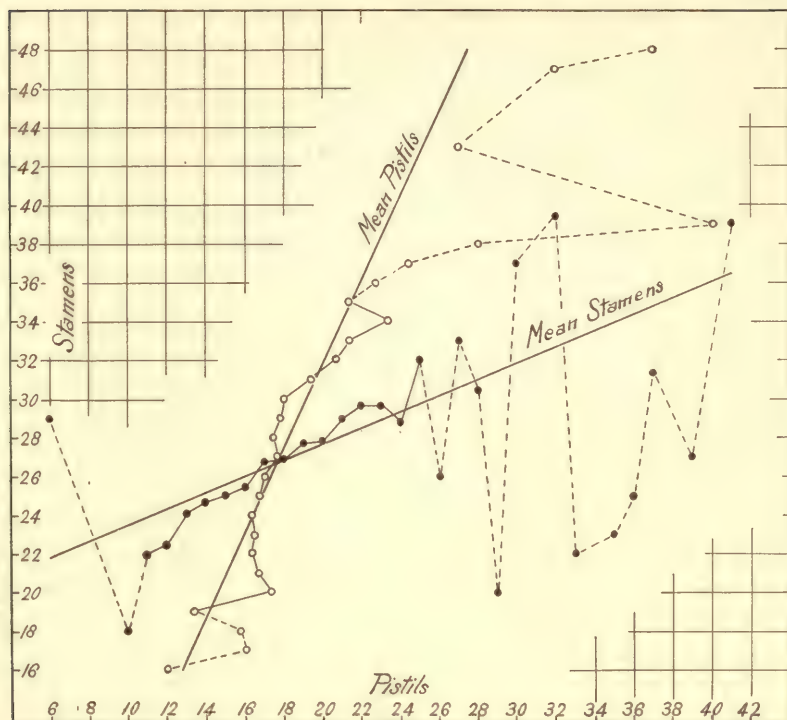
FIG. 1. Average numbers of stamens in flowers with various numbers of pistils and average numbers of pistils in flowers with various numbers of stamens. Empirical and smoothed values. Italian series.

TABLE III.

CORRELATION BETWEEN TOTAL SPOROPHYLLS AND NUMBER OF STAMENS AND PISTILS AND BETWEEN TOTAL SPOROPHYLLS AND DEVIATION OF THE NUMBER OF STAMENS AND PISTILS FROM THEIR PROBABLE VALUE.

Series.	Correlation Between Sporophylls and Stamens.	r_{sz_s}.[1]	r/E_r.	Correlation Between Sporophylls and Pistils.	r_{sz_p}.[2]
I.	.901 ± .008	− .139 ± .039	3.52	.845 ± .012	+.139 ± .039
II.	.755 ± .032	− .548 ± .053	10.37	.896 ± .015	+.548 ± .053
III.	.862 ± .011	− .378 ± .035	10.69	.873 ± .010	+.378 ± .035
IV.	.933 ± .005	− .477 ± .027	17.74	.936 ± .004	+.477 ± .027
V.	.840 ± .008	− .354 ± .024	14.99	.855 ± .007	+.354 ± .024
VI.	.836 ± .009	− .433 ± .024	18.04	.910 ± .005	+.433 ± .024
VII.	.892 ± .027	− .487 ± .023	21.37	.932 ± .004	+.487 ± .023
VIII.	.900 ± .006	− .463 ± .024	19.64	.921 ± .005	+.463 ± .024

[1] Correlation between sporophylls and deviation of stamens from their probable value.

[2] Correlation between sporophylls and deviation of pistils from their probable value.

macro- and the number of microsporophylls from their probable
value are the coefficients of critical value. These are also given
in Table III. The correlations for stamens and pistils are
necessarily equal in magnitude but opposite in sign. They
show that the number of pistils is relatively larger in the flowers
with larger numbers of sporophylls. The results are consistent
in sign throughout. All of the correlations except that for the
series from Trogen may be considered certainly significant in
comparison with their probable errors.

While the constants for certain of the series differ significantly,
the results are (considering the relatively small numbers and the
very wide geographical distribution of the material) very con-
sistent. Five of the eight series differ from $r = \pm .50$ by less
than twice their probable error. Of the other three series,
only Professor Ludwig's Trogen material is very aberrant.

For two of the series I have determined the standard devia-

Fig. 2. Empirical means and regression straight lines for regression of stamens
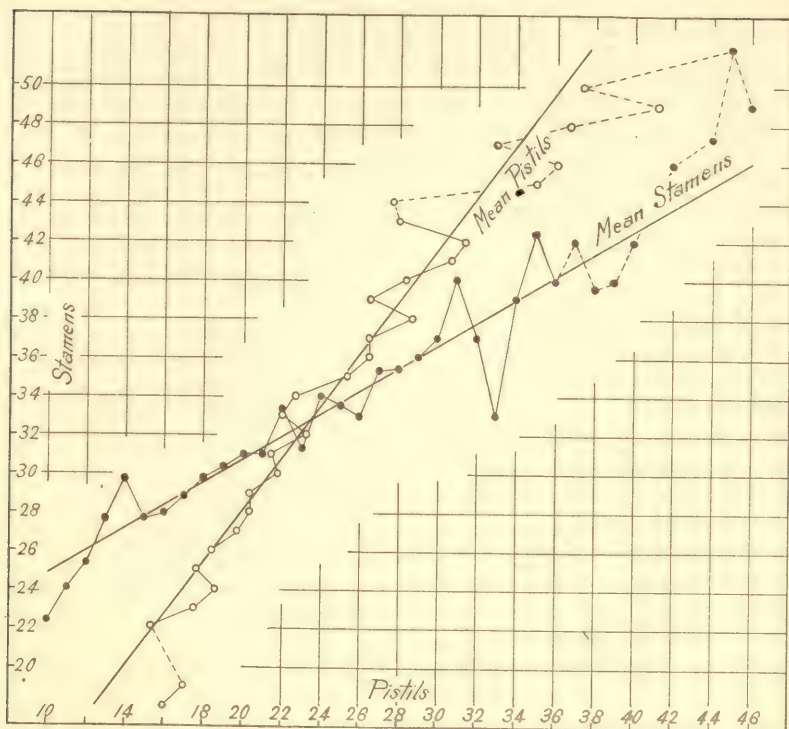on pistils and pistils on stamens. Guernsey series.

FIG. 3. Explanation as in Figs. 1 and 2. English series.

tion of the deviation of the number of stamens (or pistils) from their probable value by a formula to be published shortly ('17). They are:

For Bordighera, $\sigma_z = 2.1873$

For Surrey, $\quad \sigma_z = 2.8054$

These values make possible the determination of the straight line equations showing the regression of the deviation of the number of stamens and pistils from their probable values upon the total number of sporophylls. They are:

For Bordighera series—$Z_s = + 5.2446 - .1181\,S$

$$Z_p = - 5.2446 + .1181\,S$$

For Surrey series— $\quad Z_s = + 6.6565 - .1179\,S$

$$Z_p = - 6.6565 + .1179\,S$$

Here Z_s and Z_p denote the deviations of the stamens and pistils from their probable values and S denotes the total number of sporophylls.

The results are represented graphically in Figs. 4 and 5. In the Italian series, flowers with fewer than 36 stamens and pistils are only 29 in number, distributed among flowers with from 28–35 sporophylls per flower. The means at this end of the range cannot, therefore, be expected to show great regularity.
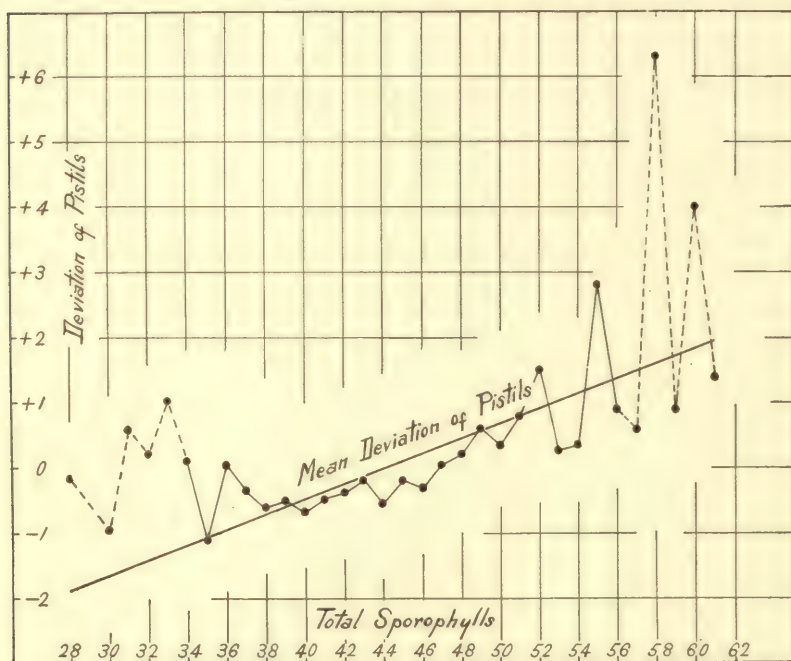

FIG. 4. Regression of the deviation of the total number of pistils from their probable value on the total number of sporophylls. Italian series.

Flowers with more than 61 sporophylls are only 10 in number but are distributed among flowers ranging from 62–85 sporophylls per flower. This portion of the range has not been included in Fig. 4. The relationship is apparently not quite linear.

In the material from Surrey, shown in Fig. 5, the increase in the relative proportion of pistils associated with increase in the total number of stamens and pistils could hardly be better represented than by the slope of the straight line indicated by the equation.
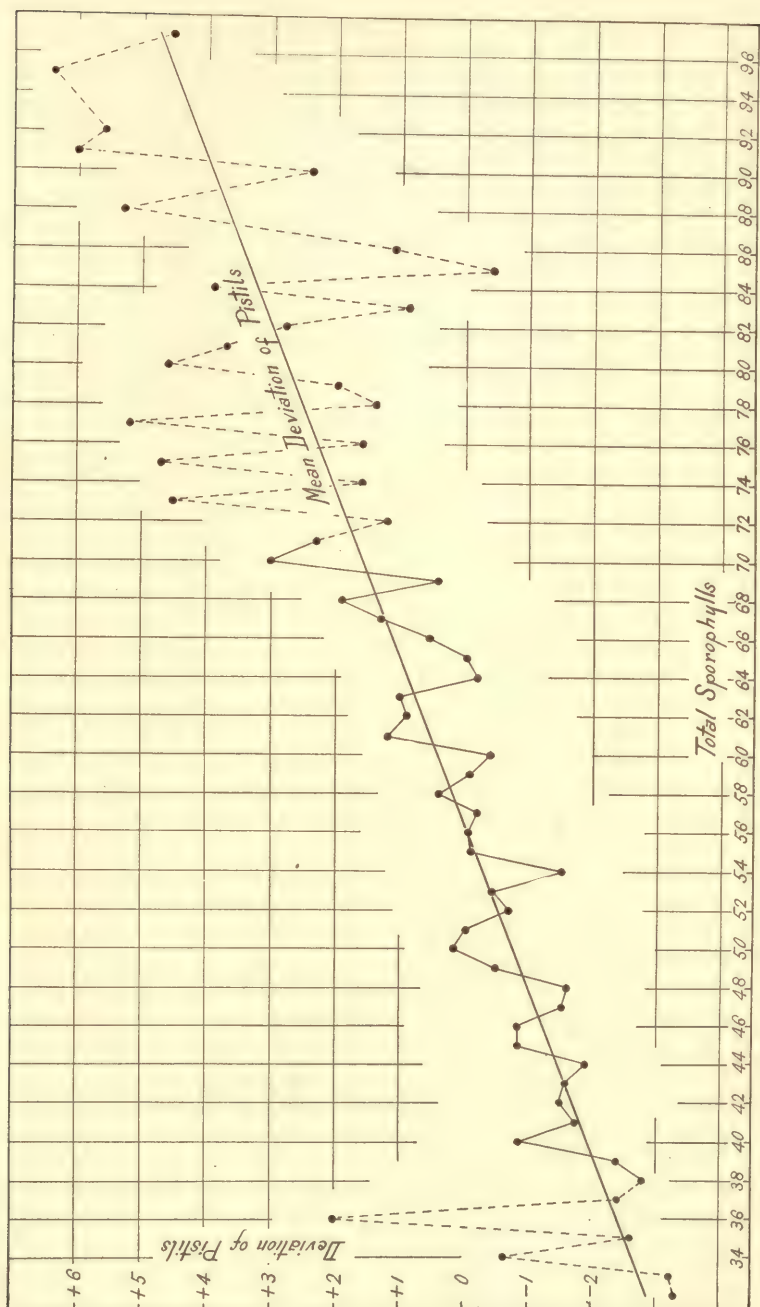
FIG. 5.   Regression of the deviation of the total number of pistils from their probable value on the total number of sporophylls. Surrey material.

IV. RECAPITULATION.

The present study deals with the problem of the relationship between the total number of sporophylls laid down and the relative number of stamens and pistils in the ranunculaceous genus *Ficaria*.

Constants have been deduced for eight series of published data from Italy, Switzerland, Belgium, Guernsey and England not hitherto analyzed by the methods now available.

In flowers with larger numbers of sporophylls the pistils are relatively more numerous than the stamens. The high degree of consistency of the results drawn from such a wide range of habitats indicates that the relationship is one of real morpho-genetic significance.

In an earlier paper it has been shown that in *Arisarum* the relative number of pistillate flowers increases as the total number of flowers becomes larger.

It is at least suggestive in relation to the problem of the physiology of sex that in both of these very different forms the number of macrosporophylls becomes relatively larger as the total number of sporophylls increases. The relationships may, however, have an embryological explanation. Only further investigations will justify final conclusions concerning the cause of the relationship demonstrated.

REFERENCES.

Harris, J. Arthur.
   '09 The Correlation between a Variable and the Deviation of a Dependent
        Variable from its Probable Value.   Biometrika, VI., pp. 438–443.
Harris, J. Arthur.
   '16 On the Distribution and Correlation of the Sexes (Staminate and Pistillate
        Flowers) in the Inflorescence of Arisarum vulgare and Arisarum probosci-
        dium.   Bull. Torr. Bot. Club, XLII., pp. 663–673.
Harris, J. Arthur.
   '17 Further Illustrations of the Applicability of a Coefficient Measuring the
        Correlation between a Variable and the Deviation of a Dependent Variable
        from its Probable Value.   Shortly.
Lee, A.
   '02 Dr. Ludwig on Variation and Correlation in *Ficaria*.   Biometrika, I.,
        pp. 316–319, 1902.
Ludwig, F.
   '01 Variationsstatistische Probleme und Materialen.   Biometrika, I., pp.
        11–29.

MacLeod, J.
  '99 Over de correlatie tusschen het aantal meeldraden en het aantal stampers bij het Speenkruid (Ficaria ranunculoides). Botanisch. Jaarboek, XI., 1899.
Pearson, K., and others.
  '03 Coöperative Investigations on Plants, II. Variation and Correlation in Lesser Celandine from Divers Localities. Biometrika, II., pp. 145–164.
Weldon, W. F. R.
  '01 Change in Organic Correlation of Ficaria ranunculoides During the Flowering Season. Biometrika, I., pp. 125–128.

THE SECOND-YEAR RECORD OF BIRDS WHICH DID AND DID NOT LAY DURING INDIVIDUAL MONTHS OF THE PULLET YEAR

The egg output of the commercial poultry plant is due in part to birds in their first and in part to birds in their second year. At some time during the first year the number of pullets is reduced to the number which is to be retained as hens during the second year.

It would be of obvious advantage if the birds sold from the flock as pullets could be those which if retained would make the poorest record in their second year. If the birds destined to be the highest producers in the second year could be selected on the basis of some criterion recognizable in the first year, it should be possible to raise the average production of the flock as a whole by increasing the average production of the hens.

In the course of a general investigation of the problem of the prediction of the egg production of the domestic fowl from the records of short periods, we have availed ourselves of the opportunity of considering the relationship between first and second year laying activity presented by the data of the Vineland International Egg Laying and Breeding Contest. As one phase of this work[1] we have sought to determine to what extent the simple criterion of *laying* versus *not laying* in any month of the first year may be used in predicting the record of the second year. The criterion has already been considered in relation to the prediction of first year egg record.[2] While our immediate purpose is the consideration of the second year production of birds which did and of those which did not lay during given months of the first year, it seems desirable to give the mean first year productions of these birds as well. For comparison the results deduced from the data of the International Egg Laying Contest at

Storrs during the year 1913-'14 and 1914-'15[3] are laid beside those presented here from the Vineland data.

The essential constants appear in the accompanying table. This gives the per cent. of the flock which did and which did not lay during the months of the first year in which any considerable proportion of the birds did not lay. The average annual production for these birds in the first year of both the Storrs (1913-'14 and 1914-'15) and the Vineland (1916-'17) contests and in the second year (1917-'18) of the Vineland contest are shown. While the actual differences in egg production are the data of practical significance, comparison between the three series is facilitated by expressing the differences between these annual means for the birds of the two classes as percentages of the actual annual average productions[4] of the flock.

Considering first the records of the pullet year we note that for the Storrs series the birds which laid in any given month show an average annual (pullet year) egg production of from 27.4 to 82.8 eggs higher than those which did not lay or from 17.9 to 54.1 per cent. For the Vineland series the difference in the production of the two groups ranges from 36.7 to 65.5 eggs or from 21.1 to 37.6 per cent. Thus the difference in the annual egg production of the birds which did and which did not lay in any given month, as well as the percentage of the birds which are not laying, varies greatly according to the month considered. During the months of November, December and January the percentage differences in the annual production of the two groups of birds is higher in the Vineland than in the Storrs series

[1] Other phases of the investigations will be reported in detail elsewhere.

[2] Harris, Blakeslee and Kirkpatrick, *Genetics*, 3: 42–44, 49–56, 1918.

[3] Harris, Blakeslee and Kirkpatrick, *loc. cit.*, p. 42.

[4] These are 153.19 eggs for the first year at Storrs, 174.05 eggs for the first year at Vineland, and 139.79 eggs for the second year at Vineland.

whereas for the other months of the eight considered the reverse is true. The average percentage difference is 30.4 in the Storrs series and 26.6 in the Vineland series.

Thus the constants show conspicuous differences of great practical significance in the first (pullet) year records of birds which did and those which did not lay during the individual months of the first year. The results for the first year records at Storrs and the first year records at Vineland are in fair agreement.

MEAN ANNUAL PRODUCTION FOR FIRST AND SECOND YEAR FOR BIRDS WHICH DID AND WHICH DID NOT LAY DURING INDIVIDUAL MONTHS OF THE FIRST YEAR

Condition of bird in month of first year	Storrs data for first year		Vineland data for first and second year		
	Per cent. of flock	First year annual mean	Per cent. of flock	First year annual mean	Second year annual mean
November					
Not laying	40.6	136.8	19.4	144.5	127.8
Laying	59.4	164.2	80.6	181.2	142.7
Difference		+ 27.4		+ 36.7	+ 14.9
		17.9%		21.1%	10.7%
December					
Not laying	38.0	133.6	22.3	142.3	125.9
Laying	62.0	165.2	77.7	183.2	143.8
Difference		+ 31.6		+ 40.9	+ 17.9
		20.6%		23.9%	12.8%
January					
Not laying	42.5	136.4	20.5	141.6	124.4
Laying	57.5	165.6	79.5	182.4	143.8
Difference		+ 29.2		+ 40.8	+ 19.4
		19.1%		23.4%	13.9%
February					
Not laying	9.9	118.6	5.0	133.6	117.0
Laying	90.1	157.0	95.0	176.2	141.0
Difference		+ 38.4		+ 42.6	+ 24.0
		25.1%		24.5%	17.2%
July					
Not laying	2.3	72.3	3.1	110.6	92.5
Laying	97.7	155.1	96.8	176.1	141.3
Difference		+ 82.8		+ 65.5	+ 48.8
		54.1%		37.6%	34.9%
August					
Not laying	5.1	89.9	7.2	121.6	99.5
Laying	94.9	156.5	92.8	178.1	142.9
Difference		+ 66.6		+ 56.5	+ 43.4
		43.5%		32.5%	31.0%
September					
Not laying	23.0	115.0	33.2	147.8	124.3
Laying	77.0	164.6	66.8	187.1	147.5
Difference		+ 49.6		+ 39.3	+ 23.2
		32.4%		22.6%	16.6%
October					
Not laying	54.7	131.9	63.2	156.9	129.8
Laying	45.3	178.9	36.8	203.5	157.0
Difference		+ 47.0		+ 46.6	+ 27.2
		30.7%		26.8%	19.5%

Turning to the second year means we note that for each of the eight months of the first year used as a basis of selection for an increase of second year preduction, the second year record of birds is higher if they laid during the special month under consideration in the first year than if they did not lay in that month. The differences between the groups amount to about two dozen eggs or more per bird in five of the eight months considered.

It is clear, therefore, that so simple a criterion as laying *vs.* non laying in the first year may furnish a criterion of some value for the selection of the birds to be retained in the flock for breeding and for second year production.

J. Arthur Harris,
Harry R. Lewis

J.A.Harris,Vol.3.